

ACGCAGACG TGCCTTCGT GATCAGGCG GCGTCGATGC TGTACGCGC ATCGGCGATG 480  
 GAGTTGCTCG AGGCGCGAGG GAACTCA 507

## (2) INFORMATION FOR SEQ ID NO:95:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 166 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

Met	Lys	Met	Val	Lys	Ser	Ile	Ala	Ala	Gly	Leu	Thr	Ala	Ala	Ala	Ala	1	5	10	15
Ile	Gly	Ala	Ala	Ala	Ala	Gly	Val	Thr	Ser	Ile	Met	Ala	Gly	Gly	Pro	20	25	30	
Val	Val	Tyr	Gln	Met	Gln	Pro	Val	Val	Phe	Gly	Ala	Pro	Leu	Pro	Leu	35	40	45	
Asp	Pro	Ala	Ser	Ala	Pro	Asp	Val	Pro	Thr	Ala	Ala	Gln	Leu	Thr	Ser	50	55	60	
Leu	Leu	Asn	Ser	Leu	Ala	Asp	Pro	Asn	Val	Ser	Phe	Ala	Asn	Lys	Gly	65	70	75	80
Ser	Leu	Val	Gln	Gly	Gly	Ile	Gly	Gly	Thr	Gln	Ala	Arg	Ile	Ala	Asp	85	90	95	
His	Lys	Leu	Lys	Lys	Ala	Ala	Gln	His	Gly	Asp	Leu	Pro	Leu	Ser	Phe	100	105	110	
Ser	Val	Thr	Asn	Ile	Gln	Pro	Ala	Ala	Ala	Gly	Ser	Ala	Thr	Ala	Asp	115	120	125	
Val	Ser	Val	Ser	Gly	Pro	Lys	Leu	Ser	Ser	Pro	Val	Thr	Gln	Asn	Val	130	135	140	
Thr	Phe	Val	Asn	Gln	Gly	Gly	Trp	Met	Leu	Ser	Arg	Ala	Ser	Ala	Met	145	150	155	160
Gln	Leu	Leu	Gln	Ala	Ala	Gly	Asn									165			

## (2) INFORMATION FOR SEQ ID NO:96:

## (1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 500 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

```

CSTGGCAATG TCGTTEACCG TCGGGGCGGG GGTGGCTTCC GCAGATCCCG TGGACGCGGT      60
CATTACACCC ACCTGCAATT ACAGGCAGGT AGTAGCTGCG CTCACGCGCA CGGATCCGGG      120
GGCTGCCGCA CAGTTCAACG CCTCAGCGGT GGCAGCTGCG TATTTGCGCA ATTCTCTGCG      180
CGCAGCGCGA CCTCAGCGCG CTGCGATGCG CGGCAATTG CAAGCTGTGC CGGGGGGCGG      240
ACAGTACATC GGCCTTCTCG AGTCGCTTGC CGGCTCTGCG AACAACTATT AAGCCCATGC      300
GGGCCCCATC CCGCGACCCG GCATCTCTCG CGGGCTTAGG CCAGATTGCG CCGCTCTCA      360
ACGGGCGCGA TCCCGCGACC CGGCATGCTC GCGGGGCTA GGCAGATTG CCCCCTCTCT      420
CAACGGGGCG CATCTCTGCG CGAATCTCTG CAGCCCGGGG GATCCACTAG TTCTAGAGCG      480
GCCGCGACCG CGGTGGAGCT                                     500
  
```

(Z) INFORMATION FOR SEQ ID NO:97:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 96 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

```

Val Ala Met Ser Leu Thr Val Gly Ala Gly Val Ala Ser Ala Asp Pro
1           5           10           15

Val Asp Ala Val Ile Asn Thr Thr Cys Asn Tyr Gly Gln Val Val Ala
20           25           30

Ala Leu Asn Ala Thr Asp Pro Gly Ala Ala Ala Gln Phe Asn Ala Ser
35           40           45

Pro Val Ala Gln Ser Tyr Leu Arg Asn Phe Leu Ala Ala Pro Pro Pro
50           55           60
  
```

130

Gln Arg Ala Ala Met Ala Ala Gln Leu Gln Ala Val Pro Gly Ala Ala  
65 70 75 80

Gln Tyr Ile Gly Leu Val Gln Ser Val Ala Gly Ser Cys Asn Asn Tyr  
85 90 95

## (2) INFORMATION FOR SEQ ID NO:98:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 134 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

ATGACAGAGC AGCAATGGAA TTCCGCGGGT ATCGAGGCCG CGGCAAGCCG AATCCAGGGA 60  
AATGTCACCT CCATTCATTC CCTCCTTGAC GAGCGGAAGC AGTCCTTGAC CAGCTCCCA 120  
GCGGCTTGGG GCGGTAGCGG TTCGGAAGCG TACC 134

## (2) INFORMATION FOR SEQ ID NO:99:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 51 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

Met Thr Glu Gln Gln Trp Asn Phe Ala Gly Ile Glu Ala Ala Ala Ser  
1 5 10 15  
Ala Ile Gln Gly Asn Val Thr Ser Ile His Ser Leu Leu Asp Glu Gly  
20 25 30  
Lys Gln Ser Leu Thr Lys Leu Ala Ala Ala Trp Gly Gly Ser Gly Ser  
35 40 45  
Glu Ala Tyr  
50

## (2) INFORMATION FOR SEQ ID NO:100:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 282 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

CGGTGCGGCA CTTCAGGTA ACTATGAAAG TCGGCTTCCG NCTGGAGGAT TCCTGAACCT	60
TCAAGCGCCG CCGATAACTG AGGTGCATCA TTAGCGGACT TTTCAGAAC ATCCTGACGC	120
GCTCGAAACG CCGCACAGCC GACGGTGGCT CCGCGAGGC GCTGCTCCA AAATCCCTGA	180
GACAATTCCG CCGGGGCGCC TACAGGAAG TCGGTCTGA ATTGCGCGG TATCTGGTCG	240
ACCTGTGTGG TGTGACCG GACCGAGCG TCCTGACCT CG	282

(2) INFORMATION FOR SEQ ID NO:101:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3058 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

GATCTATCCC GTCCAGTGC TCGGACGTT TGAGGATCGA GTGCACGTGT CTTCGTGAT	60
GGCATACCCA GAGATGTTGG CCGCGCGCGC TGACACCCCTG CAGAGCATCG GTGCTACAC	120
TGTGACTAGC AATGCCGCTG CCGCGGCCCC GACGACTGGG GTGGTGGCCC CCGCTGCCGA	180
TGAGGTGTGG GCGCTGACTG CCGCGCACTT CCGCCACAT GCGGCGATGT ATCACTCGGT	240
GAGCGCTCGG GCTGCTGCGA TTCTGACCA GTTCGTGGCC ACCCTTGGCA GCAGCGCCAG	300
CTGTATGGG GCCACTGAAG TCGCCAATGC GCGCGCGCCC ACCTAAGCGA GGAACAGTCG	360
GCACGAGAAA CCACGACAAA TAGGGACAGC TAATGCTTGA TTTCGGGGGG TTACACCGGG	420
AGATCAACTC CCGGAGGATG TAGCGCGGCC CGGCTTGGC CTCGCTGGTG GCGCGGCTC	480
AGATGTGGCA CAGCTGTCG AGTGACCTGT TTTCGGCCGC GTGGCGCTTT CAGTCGCTGG	540
TCYGGGCTCT GACGCTGGGG TCGTGATAG GTTCGTGGCC GGGTCTGATG GTGGCGGGCG	600

CCTCGCCGTA	TGTGGCTGG	ATGAGGCTCA	CCGCGGGGCA	GGCCGAGCTG	ACCGCCGCCC	660
AGGTCCGGGT	TGCTCGGGCG	GCTACGAGA	CGGATATGG	GCTGACGGTG	CCCCCGCCCG	720
TCATCGCCGA	GAACGCTGCT	GAAGTGATGA	TTCTGATAGC	GACCAACCTC	TTGGGGCAAA	780
ACACCCCGGC	GATCGCGGTC	AACGAGGCGG	AATACGGCGA	GATGTVGGCC	CAAGACGCGG	840
CCGCGATGTT	TGGCTACGCC	GCGGCGACGG	CGACGGCGAC	GGCGAGCTTG	CTGCCGTTGG	900
AGGAGGCGCC	GGAGATGACC	AGCGGGGGTG	GGCTCTCTGA	GCAGGCGCGC	GCGGTGAGGG	960
AGGCCTCCGA	CACCGCGCGG	GCGAACCAAT	TGATGAACAA	TGTGCCCCAG	GCGCTGCANC	1020
AGCTGGCCCA	GCCACGCGAG	GGCACCACGC	CTTCTTCCAA	GCTGGCTGGC	CTGTGAGAGA	1080
CGGTCTCGCC	GCATCGGCTG	CGATCAGCA	ACATGGTGTG	GATGCCCAAC	AACCACATGT	1140
CGATGACCAA	CTCGGGTGTG	TGATGACCA	ACACCTTGAG	CTCGATGTTG	AAGGGCTTTG	1200
CTCCGGCGGC	GGCGGCGCAG	GCKGTGCAAA	CCGCGCGGCA	AAACGGGGTC	CGGGCGATGA	1260
GCTCGCTGGG	CAGCTCGCTG	GGTTCCTCGG	GTCTGGGCGG	TGGGTTGGCC	GCCAACTTGG	1320
GTGCGGCGGC	CTCGGTCCGT	TGGTTCTCGG	TGCGGCAGGC	CTGGCGCGCG	GCCAAACGAG	1380
CAGTCACCCC	GGCGGCGCGG	GCGCTGCGCG	TGACCAAGCT	GACCAAGCGC	GCGGAAAGAG	1440
GGCCCGGGCA	GATGCTGGGC	GGGCTGCGCG	TGGGGCAGAT	GGCGGCGAGG	GGCGGTGGTG	1500
GGCTCACTGG	TGTGCTCGGT	GTTCGCGCGC	GACCTATGT	GATGCGCGAT	TCTCGCGCGG	1560
CCGGCTAGGA	GAGGGGGCGG	AGACTGTGCT	TATTTGACCA	GTCATCGCGG	GTCTCGGTGT	1620
TTCCGCGCGC	GGCTATGACA	ACAGTCAATG	TGATGACAA	GTCACAGGTA	TTAGGTCCAG	1680
GTTCAACAGG	GAGACAGGCA	ACATGGCCTC	AGGTTTATG	ACGGATCCGC	ACGCGATCCG	1740
GGACATGGCG	GGCGTFTTTG	AGGTGCACGC	CCAGACGGTG	GAGGACGAGG	CTCGCCGGAT	1800
GTGGGCGTCC	GCGCAAAACA	TTTCCGCTGC	GGGCTGGAGT	GGCATGCGCG	AGGGGACCTC	1860
GCTAGACACC	ATGCCCCAGA	TGAATCAGGC	GTTCGCAAC	ATCGTGAACA	TGCTGCACCG	1920
GGTGCCTGAC	GGGCTGCTTC	GCGACGCCAA	CAACTACGAG	CAGCAAGAGC	AGGCTTCCCA	1980
GCAGATCCTC	AGCAGCTAAC	GTCAGCGCTT	GCAGCACAA	ACTTTTACAA	GCGAAGGAGA	2040
ACAGGTTCCA	TGACCATCAA	CTATCAATTC	GGGATGTGG	ACGCTCACGG	CGCATGATTC	2100
CGCGCTCAGG	CCGGTTTGGT	GGAGGCGGAG	CATCAGGCCA	TCATTCGTGA	TGTGTTGAGC	2160
GCGAGTGAAT	TTTGGGGCGG	CGCGGCTTGG	GCGGCTGCGC	AGGGGTTGAT	TAGCCAGTTG	2220
GGCGTAAGT	TCCAGCTGAT	CTAGGAGCAG	GCCAAACGCC	ACGGGCGAGAA	GCTGCGAGCT	2280

```

GCCGGCAACA ACATGGGCGA AACCGACAGC GCCGTGGGCT CCAGCTGGGC CTGACACCAG 2340
GCCAAGGCCA GGGACGTGGT GTACGAGTGA AGTTCTCTGC GTGATCCTTC GGGTGCCAGT 2400
CTAAGTGGTC AGTGCTGGGG TGTTGGTGGT TTGCTGCTTG CGGGGTTCTT CGGTGCTGGT 2460
CAGTGCTGCT CGGGCTCGGG TGAGGACCTC GAGGCCDAGG TAGCGCCGTC CTTCGATCCA 2520
TTGCTGCTGT TGTTCGGCGA GGACGGCTCC GACGAGGCGG ATGATCGAGG CGCGGTCCGG 2580
GAAGATGCCC ACCACGTCCG TTGGGCGTGG TACCTGTCCG TTGAGGCGTT CTTGGGGGTT 2640
GTTGAGCCAG ATTTGGCGCC AGATCTGCTT GGGGAAGGCG GTGAACGCCA CGAGGTGGGT 2700
GCGGGCGGTC TGAGGTGCT CCGCCACCGC GGGCAGTTG TCGGTGAGAG CGTCGAGTAC 2760
CCGATCATAT TGGGCAACAA CTGATTCCGC GTCGGGCTGG TCGTAGATGC ACTGCAGCAG 2820
GCTCGGCACC CACCGCCAGG AGGCTTGGG GGTGGCTGCC ATCAGATTGG CTGCGTAGTG 2880
GGTTCTGCG CGCTGCCAGG CCGCTGCGGG CAGGCTGGCG CCGATCGCGG CCACCAGGCC 2940
GGGCTGGGCG TGCTGTGTGA CCAGCGCGAC CCGGACAGC CCGCGGGGCA CCAGGTCCGG 3000
GAAGAACGCG AGCCAGCCGG CCGGCTCCTC GCGCGAGGTC ACCTGGATGC CCAGGATC 3058

```

## (2) INFORMATION FOR SEQ ID NO:102:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 391 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

```

Met Val Asp Phe Gly Ala Leu Pro Pro Glu Ile Asn Ser Ala Arg Met
1           5           10          15
Tyr Ala Gly Pro Gly Ser Ala Ser Leu Val Ala Ala Ala Gln Met Trp
20          25          30
Asp Ser Val Ala Ser Asp Leu Phe Ser Ala Ala Ser Ala Phe Gln Ser
35          40          45
Val Val Trp Gly Leu Thr Val Gly Ser Trp Ile Gly Ser Ser Ala Gly
50          55          60
Leu Met Val Ala Ala Ala Ser Pro Tyr Val Ala Trp Met Ser Val Thr
65          70          75          80

```

134

Ala	Gly	Gln	Ala	Glu	Leu	Thr	Ala	Ala	Gln	Val	Arg	Val	Ala	Ala	Ala	
				85					90					95		
Ala	Tyr	Glu	Thr	Ala	Tyr	Gly	Leu	Thr	Val	Pro	Pro	Pro	Val	Ile	Ala	
				100				105					110			
Glu	Asn	Arg	Ala	Glu	Leu	Met	Ile	Leu	Ile	Ala	Thr	Asn	Leu	Leu	Gly	
				115			120					125				
Gln	Asn	Thr	Pro	Ala	Ile	Ala	Val	Asn	Glu	Ala	Gln	Tyr	Gly	Glu	Met	
				130		135					140					
Trp	Ala	Gln	Asp	Ala	Ala	Ala	Met	Phe	Gly	Tyr	Ala	Ala	Ala	Thr	Ala	
145					150				155						160	
Thr	Ala	Thr	Ala	Thr	Leu	Leu	Pro	Phe	Gln	Gln	Ala	Pro	Glu	Met	Thr	
				165				170						175		
Ser	Ala	Gly	Gly	Leu	Leu	Glu	Gln	Ala	Ala	Ala	Val	Glu	Glu	Ala	Ser	
				180				185					190			
Asp	Thr	Ala	Ala	Ala	Asn	Gln	Leu	Met	Asn	Asn	Val	Pro	Gln	Ala	Leu	
				195			200					205				
Gln	Gln	Leu	Ala	Gln	Pro	Thr	Gln	Gly	Thr	Thr	Pro	Ser	Ser	Lys	Leu	
				210		215					220					
Gly	Gly	Leu	Trp	Lys	Thr	Val	Ser	Pro	His	Arg	Ser	Pro	Ile	Ser	Asn	
225					230				235						240	
Met	Val	Ser	Met	Ala	Asn	Asn	His	Met	Ser	Met	Thr	Asn	Ser	Gly	Val	
				245				250						255		
Ser	Met	Thr	Asn	Thr	Leu	Ser	Ser	Met	Leu	Lys	Gly	Phe	Ala	Pro	Ala	
				260				265				270				
Ala	Ala	Ala	Gln	Ala	Val	Gln	Thr	Ala	Ala	Gln	Asn	Gly	Val	Arg	Ala	
				275			280				285					
Met	Ser	Ser	Leu	Gly	Ser	Ser	Leu	Gly	Ser	Ser	Gly	Leu	Gly	Gly	Gly	
				290		295					300					
Val	Ala	Ala	Asn	Leu	Gly	Arg	Ala	Ala	Ser	Val	Gly	Ser	Leu	Ser	Val	
305					310					315					320	
Pro	Gln	Ala	Trp	Ala	Ala	Ala	Asn	Gln	Ala	Val	Thr	Pro	Ala	Ala	Arg	
				325				330						335		
Ala	Leu	Pro	Leu	Thr	Ser	Leu	Thr	Ser	Ala	Ala	Glu	Arg	Gly	Pro	Gly	
				340				345					350			
Gln	Met	Leu	Gly	Gly	Leu	Pro	Val	Gly	Gln	Met	Gly	Ala	Arg	Ala	Gly	
				355			360					365				

Gly Gly Leu Ser Gly Val Leu Arg Val Pro Pro Arg Pro Tyr Val Met  
 370 375 380

Pro His Ser Pro Ala Ala Gly  
 385 390

## (2) INFORMATION FOR SEQ ID NO:103:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1725 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

GACGTCAAGCA	CCCCCCCTGC	AGGGCTGGAG	CGTGGTGGT	TTTCATCTGC	GGTCAAGGTC	60
ACGTCCCTTC	GCGTGTGCCC	GGGTGGGATG	CAGACTCGAT	CCCCCTCTTT	AGTGCAGACTA	120
ATTTCGTTCA	AGTCCCTGCC	AGGTATAGGA	CTTCACGATC	GGTTAATGTA	GGGTTCACCC	180
CGTGTGGGG	TGGATTGGC	CGGACCAATC	GTCAACCAAG	CTTGGCGTGC	GGGCCAGGCG	240
GGCGATCAGA	TGCTTCACT	ACCAATCAAT	CTTGAGCTGC	CGGGCCGATG	CTCGGGCTAA	300
ATGAGGAGGA	GCACGGGTGT	CTTCACTGC	GCAACCGGAG	ATGTTGGCGG	CCGCGGCTGG	360
CGAACTTCGT	TCCCTGGGG	CAACGCTGAA	GGCTAGCAAT	GGGCCCCAG	CGGTGCGGAC	420
GACTGGGGTG	GTGCCCCGG	CTGCGGAGGA	GGTGTGCGTG	CTGCTTGCCA	CACAAITCCG	480
TACCATGCG	GCGACGTATC	AGACGGCCAG	CGCCAAGGCC	CGGGTGATCC	ATGAGCAGTT	540
TGTGACCACG	CTGGCCACCA	GCGCTAGTTT	ATATGCGGAC	ACCGAGGCCG	CCACCGCTGT	600
GGTCACCGGC	TAGCTGACCT	GACGGTATTC	GAGCGGAAGG	ATTATCGAAG	TGGTGCATTT	660
CGGGGCGTTA	CCACCGGAGA	TCAACTCCGC	GAGGATGTAC	GGGGCCCCGG	GTTCGGGCTC	720
GCTGTTGGCC	GCCGCGAAGA	TGTGGGACAG	CGTGGCGAGT	GACCTGTTTT	CGGCCGCTC	780
GGCGTTTCAG	TGGTGGTCT	GGGTCTGAC	GGTGGGCTCG	TGGATAGGTT	CGTGGGCGGG	840
TCTGATGGCG	CGGGCGGCGT	CGCGGTATGT	GGGTGGGATG	AGCGTCAGCG	CGGGGCGGGC	900
CCACCTGACC	GCGGCGCAGG	TGGGGGTTCC	TGGGGCGGCC	TACGAGACAG	CGTATAGGCT	960
BACGGTGCCC	CGGCGGCTGA	TGGCGGAGAA	CGGTACCGAA	CTGATGACGC	TGACCGCGAC	1020
CAACCTCTTC	GGGCAAAACA	CGCGGGCGAT	CGAGGCCAAT	CAGGCGGCAT	ACAGCCAGAT	1080



```

GTGGGGCCAA  GACGCGGAGG  CGATGTATGG  CTACGCCGCC  ACCGCGGCGA  CGGCGACCGA  1140
GGCGTTGCTG  CCGTTCCAGG  ACCGCCCACT  GATCAGCAAC  CCGGGCGGGC  TCGTTGAGCA  1200
GGCGTCCGCG  GTCCAGGAGG  CCATCGACAC  CGCGCGGCGG  AACCACTTGA  TGAACAATGT  1260
GCCCCAAGCG  CTGCAACAGC  TGGGCGAGCC  AGCGCAGGGC  GTCGTACCTT  CTTCCAGAGT  1320
GGGTGGGCTG  TGGACGGGGC  TCTCGCGGCA  TCTGTCCGCG  CTCAGCAACG  TCACTTCGAT  1380
AGCCAACAAC  CACATGTCCA  TGATGGGCAC  GCGTGTGTGG  ATGACCAACA  CCTTGCACTC  1440
GATGTTGAAG  GCCTTACTTC  CGCGCGCGGC  TCAGGCCGTG  GAAACCGCGG  CGCAAAACGG  1500
GGTCTGGGCG  ATGAGCTGGC  TGGGCGAGCA  GCTGGGTTGG  TCGCTGGGTT  CTTGGGTTCT  1560
GGGCGCTGGG  GTGGCGGCGA  ACTTGGGTCG  GCGGCGCTCG  GTGGGTTGCT  TGTGGGTGCC  1620
GCCAGCATGG  GCGCGGGCGA  ACCAGCGGGT  CACCGCGGCG  GCGCGGGCGC  TCGCGGTGAC  1680
CAGCCTGACC  AGCGCGCGCC  AAGTCGCCCC  CGGACACATG  CTGGG

```

## (2) INFORMATION FOR SEQ ID NO:104:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 359 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) SEQUENCE DESCRIPTION: SEQ ID NO:104:

```

Val Val Asp Phe Gly Ala Leu Pro Pro Glu Ile Asn Ser Ala Arg Met
1           5           10          15
Tyr Ala Gly Pro Gly Ser Ala Ser Leu Val Ala Ala Ala Lys Met Trp
20          25          30
Asp Ser Val Ala Ser Asp Leu Phe Ser Ala Ala Ser Ala Phe Gln Ser
35          40          45
Val Val Trp Gly Leu Thr Val Gly Ser Trp Ile Gly Ser Ser Ala Gly
50          55          60
Leu Met Ala Ala Ala Ala Ser Pro Tyr Val Ala Trp Met Ser Val Thr
65          70          75          80
Ala Gly Gln Ala Gln Leu Thr Ala Ala Gln Val Arg Val Ala Ala Ala
85          90          95

```

Ala Tyr Glu Thr Ala Tyr Arg Leu Thr Val Pro Pro Pro Val Ile Ala  
 100 105 110  
 Glu Asn Arg Thr Glu Leu Met Thr Leu Thr Ala Thr Asn Leu Leu Gly  
 115 120 125  
 Glu Asn Thr Pro Ala Ile Glu Ala Asn Gln Ala Ala Tyr Ser Gln Met  
 130 135 140  
 Trp Gly Gln Asp Ala Glu Ala Met Tyr Gly Tyr Ala Ala Thr Ala Ala  
 145 150 155 160  
 Thr Ala Thr Glu Ala Leu Leu Pro Phe Glu Asp Ala Pro Leu Ile Thr  
 165 170 175  
 Asn Pro Gly Gly Leu Leu Glu Gln Ala Val Ala Val Glu Glu Ala Ile  
 180 185 190  
 Asp Thr Ala Ala Ala Asn Gln Leu Met Asn Asn Val Pro Gln Ala Leu  
 195 200 205  
 Glu Gln Leu Ala Gln Pro Ala Gln Gly Val Val Pro Ser Ser Lys Leu  
 210 215 220  
 Gly Gly Leu Trp Thr Ala Val Ser Pro His Leu Ser Pro Leu Ser Asp  
 225 230 235 240  
 Val Ser Ser Ile Ala Asn Asn His Met Ser Met Met Gly Thr Gly Val  
 245 250 255  
 Ser Met Thr Asn Thr Leu His Ser Met Leu Lys Gly Leu Ala Pro Ala  
 260 265 270  
 Ala Ala Gln Ala Val Glu Thr Ala Ala Gln Asn Gly Val Trp Ala Met  
 275 280 285  
 Ser Ser Leu Gly Ser Gln Leu Gly Ser Ser Leu Gly Ser Ser Gly Leu  
 290 295 300  
 Gly Ala Gly Val Ala Ala Asn Leu Gly Arg Ala Ala Ser Val Gly Ser  
 305 310 315 320  
 Leu Ser Val Pro Pro Ala Trp Ala Ala Ala Asn Gln Ala Val Thr Pro  
 325 330 335  
 Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr Ser Ala Ala Gln Thr  
 340 345 350  
 Ala Pro Gly His Met Leu Gly  
 355

(2) INFORMATION FOR SEQ ID NO:105;

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3027 base pairs

(B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:185:

AGTTCAGTCC	AGAATGATAC	TCACGGGCTG	TATCCACGAT	GGCTGAGACA	ACCGAACCAC	60
CGTGGGACGC	GGGGACATCG	CAAGCCGACG	CGATGGCGTT	GGCGGCGGAA	GCCGAAGCCG	120
CCGAAGCCGA	AGCGCTGGCC	GCGCGGCGCC	GGCGCGCTGC	CCGTGCGGCC	CGTTGAGGC	180
GTGAGGCGCT	GGCGATGGCC	CCAGCCGAGG	ACGAGAACGT	CCCCGAGGAT	ATGCAGACTG	240
GGAAGACGCC	GAAGACTATG	ACGACTATGA	CGACTATGAG	GCCGAGAGCC	AGGAGGCGCC	300
ACGTTGCGCA	TCCTGGCGAC	GGCGCTTCCG	GCTGCGGTTA	CCAAGACTGT	CCAGGATTGC	360
CATGGCGGCC	GCAGTCTGCA	TCATCTGCGG	CTTCACCGGG	CTCAGCGGAT	ACATTGTGTG	420
GCACACCCAT	GAGGCCACCG	AACGCCAGCA	GCGCGCGCGG	GCGTTCGCGG	CCGGAOCCAA	480
GCAAGCTGTC	ATCAACATGA	CCTCGCTGGA	CTTCAACAAG	GCCAAAGAAG	ACGTGCGCGG	540
TGTGATCGAC	AGCTOCACCG	GCGAATTCAG	GGATGACTTC	CAGCAGCGGG	CAGCCGATTT	600
CACCAAGGTT	GTGGAACGAT	CCAAAGTGGT	CACCGAAGGC	ACGGTGAAAG	CGACAGCCGT	660
CGAATCCATG	AACGAGCAAT	CCGCGGTGGT	GCTGCTGCGG	GCGACTTCAC	GGTTCACCAA	720
TTGCGGTGGG	GCGAAAGAGG	AACCACTGCG	GTGCGGCGTC	AAAGTGACCG	TCACCGAAGA	780
GGGCGGACAG	TACAGATGT	CGAAGTTGA	GTTCGTTACG	TGACCGATGA	CGTACGCGAC	840
GTCAACACCG	AAACCACTGA	CGCCACCGAA	GTGCTGAGA	TGACTCAGC	CGCAGGCGAA	900
GCGCGTGATT	GCGCGACCGA	GCCATTGAC	ACCGACTCTG	CAACCGAATC	TACCGGCGAG	960
AAGGCTCAGC	GCGACCGTGA	CCTGTGGCGA	ATGCAGGTTA	CCTTGAAACC	CGTTCCGGTG	1020
ATTCTCATCC	TGCTCATGT	GATCTCTGGG	GGCGCGACCG	GATGCTATA	CCTTGAGCAA	1080
TACGACCCGA	TCAGCAGAGG	GACTCGGGCG	CGCGCGGTGC	TGCGGTGCGC	GCGGCGCTCG	1140
ACGGGACAAAT	GCGCTGTTC	TGTATTGACC	CGACAGCTCG	ACCAAGACTT	CGCTACCGCC	1200
AGGTGCGACC	TGCGCGCGGA	TTTCTGTGCG	TATACGACCA	GTTACGCGAG	CAGATCGTGG	1260
CTCGGCGCGC	CRAACAGAGG	TCACTGAAAA	CCACCGCCAA	GCTGCTGCGC	GCGGCGGTGT	1320
CGGAGCTACA	TCCGGATTCC	GCGCTGCTTC	TGCTTTTTGT	CGACCAAGAG	ACTACCAATA	1380

AGGACAGCCC	CAATCCCTCG	ATGGCGGCCA	GCAGCCTEAT	GGTGACCCTA	GCCAAGGCTG	1440
ACGGCAATTC	GCTGATCACC	AAGTTCACCC	CGGTTTAGGT	TGCGGTAGGC	GGTCGCCAAG	1500
TCTAGCGGGG	GCAGCGGTGG	CTGCTCGTGC	GAGATACCGG	CCCTTCCTCG	GACAATCAGC	1560
GCCCCACCTC	AAACAGATCT	CGGCCCTGT	CTAATCGGCC	GGTTATTTA	AGATTAGTTG	1620
CCACTGTATT	TACCTGATGT	TCAGATTGTT	CAGCTGATT	TAGCTTCGGG	GCAGGGCGGC	1680
TGGTGCACTT	TGCATCTGGG	GTTGTGACTA	CTTGAGAGAA	TTTGACCTGT	TGCCGACGTT	1740
GTTTCTCTGC	CATCATTCGT	GCTAGTTATG	GCCGAGCGGA	AGGATTAATG	AAGTGGTGGG	1800
CTTCGGGGCG	TTACCAACCG	AGATCAACTC	CCGAGGATG	TACGCCCGCC	CGGTTTCGGC	1860
CTCGCTGGTG	GCGCGCGCCA	AGATCTGGGA	CACCGTGGCG	AGTGACCTGT	TTTCGGGGCG	1920
CTCGGCCTTT	CAGTCGGTGG	TCTGGGCTCT	GACGACGGGA	TGGTGATAG	GTTGGTCCGC	1980
GGGTCTGATG	GTGGCGGCGG	CCTCGCGGTA	TGTGGCGTGG	ATGAGCGTCA	CCGCGGGGCA	2040
GGCGGAGCTG	ACCGCTGCCC	AGGTCCGGGT	TGCTGCGGGG	GCTTACGAGA	CGCGGTATGG	2100
GCTGACGGTG	CCCCCGCCCC	TGATCGCCGA	GAACCGTGCT	GAAGTGATCA	TTCTGATAGC	2160
GACCAACCTC	TTGGGGCAAA	ACACCCCGGC	GATCGCGGTC	AACGAGGCGG	AATACGGGGA	2220
GATGTGGGCC	CAAGACGCCG	CCGCGATGTT	TGGCTAGGTC	GCCACGGCGG	CGACGGCGAC	2280
CGAGGCGTTG	CTGCCGTTG	AGGACGCTCT	ACTGATCACU	AACCCCGGCG	GGCTCCTTGA	2340
GCAGGCGCTC	GCGGTGAGGG	AGGCCATCGA	CACCGCCGCG	GCGAACCGGT	TCATGAACAA	2400
TGTGCCCCAA	GCCTGCAAC	AAC TGSCCA	GCCACGAAA	AGCATCTGGC	CGTTCGACCA	2460
ACTGAGTGAA	CTCTGGAAG	CCATCTCGCC	GCATCTGTGG	CCGCTCAGCA	ACATCGGTTC	2520
GATGCTCAAC	AACCAAGTGT	CGATGACCAA	CTCGGTTGTG	TGGATGCGCA	GCACCTTGCA	2580
CTCAATGTTG	AAGGGCTTTG	CTCGGGCGGC	GGCTCAGGCC	GTGGAAACCG	CGCGGCAAAA	2640
CGGGGTCCAG	GCGATGAGCT	CGCTGGGCA	CCAGCTGGGT	TGCTCGCTGG	GTTCTTCGGG	2700
TCTGGGGGCT	GGGGTGGCGG	CCAACTTGGG	TGGGGCGGCC	TGGTTCGGTT	CGTTGTGGGT	2760
GCGGCAGGCC	TGGGCGCGCG	CCAACGAGGC	GCTCACCCCG	GCGGCGCGGG	CGCTGCGGCT	2820
GACCAAGCTG	ACCAGCGCGG	CCCAAACCGC	CCCCGGACAC	ATGCTGGGCG	GGCTACGGCT	2880
GGGCAACTG	ACCAATAGCG	GCGGCGGGTT	CGGCGGGGTT	AGCAATGCGT	TGGGGATGCC	2940
GCGCGGGGCG	TACGTAATGC	CCCGTGTGCC	CGCGCGCGGG	TACGCGCGAT	CCGCACGCAA	3000

TCCGGGCCCC CTATGCCCC AGCGATC

3027

## (2) INFORMATION FOR SEQ ID NO:106:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 396 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

Val	Val	Asp	Phe	Gly	Ala	Leu	Pro	Pro	Glu	Ile	Asn	Ser	Ala	Arg	Met	1	5	10	15
Tyr	Ala	Gly	Pro	Gly	Ser	Ala	Ser	Leu	Val	Ala	Ala	Ala	Lys	Met	Trp	20	25	30	
Asp	Ser	Val	Ala	Ser	Asp	Leu	Phe	Ser	Ala	Ala	Ser	Ala	Phe	Gln	Ser	35	40	45	
Val	Val	Trp	Gly	Leu	Thr	Thr	Gly	Ser	Trp	Ile	Gly	Ser	Ser	Ala	Gly	50	55	60	
Leu	Met	Val	Ala	Ala	Ala	Ser	Pro	Tyr	Val	Ala	Trp	Met	Ser	Val	Thr	65	70	75	80
Ala	Gly	Gln	Ala	Glu	Leu	Thr	Ala	Ala	Gln	Val	Arg	Val	Ala	Ala	Ala	85	90	95	
Ala	Tyr	Glu	Thr	Ala	Tyr	Gly	Leu	Thr	Val	Pro	Pro	Pro	Val	Ile	Ala	100	105	110	
Glu	Asn	Arg	Ala	Glu	Leu	Met	Ile	Leu	Ile	Ala	Thr	Asn	Leu	Leu	Gly	115	120	125	
Gln	Asn	Thr	Pro	Ala	Ile	Ala	Val	Asn	Glu	Ala	Glu	Tyr	Gly	Glu	Met	130	135	140	
Trp	Ala	Gln	Asp	Ala	Ala	Ala	Met	Phe	Gly	Tyr	Ala	Ala	Thr	Ala	Ala	145	150	155	160
Thr	Ala	Thr	Glu	Ala	Leu	Leu	Pro	Phe	Glu	Asp	Ala	Pro	Leu	Ile	Thr	165	170	175	
Asn	Pro	Gly	Gly	Leu	Leu	Glu	Gln	Ala	Val	Ala	Val	Glu	Glu	Ala	Ile	180	185	190	
Asp	Thr	Ala	Ala	Ala	Asn	Gln	Leu	Met	Asn	Asn	Val	Pro	Gln	Ala	Leu	195	200	205	

Gln	Gln	Leu	Ala	Gln	Pro	Thr	Lys	Ser	Ile	Trp	Pro	Phe	Asp	Gln	Leu
210							215				220				
Ser	Gln	Leu	Trp	Lys	Ala	Ile	Ser	Pro	His	Leu	Ser	Pro	Leu	Ser	Asn
225					230					235					240
Ile	Val	Ser	Met	Leu	Asn	Asn	His	Val	Ser	Met	Thr	Asn	Ser	Gly	Val
				245						250				255	
Ser	Met	Ala	Ser	Thr	Leu	His	Ser	Met	Leu	Lys	Gly	Phe	Ala	Pro	Ala
			260					265					270		
Ala	Ala	Gln	Ala	Val	Glu	Thr	Ala	Ala	Gln	Asn	Gly	Val	Gln	Ala	Met
		275					280					285			
Ser	Ser	Leu	Gly	Ser	Gln	Leu	Gly	Ser	Ser	Leu	Gly	Ser	Ser	Gly	Leu
	290					295					300				
Gly	Ala	Gly	Val	Ala	Ala	Asn	Leu	Gly	Arg	Ala	Ala	Ser	Val	Gly	Ser
305					310					315					320
Leu	Ser	Val	Pro	Gln	Ala	Trp	Ala	Ala	Ala	Asn	Gln	Ala	Val	Thr	Pro
				325					330					335	
Ala	Ala	Arg	Ala	Leu	Pro	Leu	Thr	Ser	Leu	Thr	Ser	Ala	Ala	Gln	Thr
		340						345					350		
Ala	Pro	Gly	His	Met	Leu	Gly	Gly	Leu	Pro	Leu	Gly	Gln	Leu	Thr	Asn
		355					360					365			
Ser	Gly	Gly	Gly	Phe	Gly	Gly	Val	Ser	Asn	Ala	Leu	Arg	Met	Pro	Pro
	370					375					380				
Arg	Ala	Tyr	Val	Met	Pro	Arg	Val	Pro	Ala	Ala	Gly				
385					390					395					

## (2) INFORMATION FOR SEO ID NO: 107:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1616 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(a) SEQUENCE DESCRIPTION: SEQ ID NO:107:

CATCGGAGGG	ASTGATCACC	ATGCTGTGGC	ACUAAATGCC	ACCGGAGTAA	ATGCCGCACC	60
GCTGATGGCC	GGCGGGGGTC	CGGCTCCAAAT	GCTTGGGGCS	GGGGGGGAT	GGCAGAGGCT	120
TTTGGGGGCT	CTGAGCGGTC	AGGGGTGGA	GTTCAGGGCS	CGGCTGAACT	CTCTGGGAGA	180

AGCCTGGACT	GGAGGTGGCA	CCGACAAGGC	GCTTCCGGCT	GCAACGCCGA	TGGTGGTCTG	240
GCTACAAACC	GGTCAACAC	AGGCCAAGAC	CCGTGGGATG	CAGGCGACGG	CGCAAGCCGC	300
GGCATAACCC	CAGGCCATGG	CCACGACGCC	GTCGCTGCGG	GAGATCGCCG	CCAACCACAT	360
CACCCAGGCC	GTCTTACGG	CCACCACCTT	CTTCGGTATC	AACACGATCC	CGATCGGCTT	420
GACCGAGATG	GATTATTTCA	TGGTATGTG	GAACCAGGCA	GCCCTGGCAA	TGGAGGTCTA	480
CCAGGCCGAG	ACCCGCTTA	ACACGCTTTT	CGAGAAGCTC	GAGCCGATGG	CGTCGATCCT	540
TGATCCCGGC	GCGAGCCAGA	GCACGACGAA	CCCGATCTTC	GGAATGCCCT	CCCTTGGCAG	600
CTCACACCGG	GTTCGCCAET	TGCCGCGGBC	GGCTACCCAG	ACCCTCGGCC	AACTGGGTGA	660
GATGAGCGGC	CUGATGCAGC	AGCTGACCCA	GCCTCTCCAG	CAGGTGACCT	CGTTCTTCAG	720
CCAGGTGGGC	GGACCGGCG	GCAGCAACCC	AGCCGACGAG	GAGGCCGCCG	AGATGGGCCCT	780
GCTCGGCACC	AGTCGCTGT	CGAACCTCC	GCTGCTGGT	CGATCAGGCC	CCAGCCGCGG	840
CGCGGGCCTG	CTGCGCGCGG	ASTGCTTACC	TGGCGCAGGT	GGTGGTTGA	CCCGCACGCC	900
GCTGATGTCT	CAGCTGATCG	AAAAGCCGGT	TGCCCCCTCG	CTGATGCCCG	CGGCTGCTCC	960
CGGATCGTGG	GCGACGGGTG	GCGCGGCTCC	GGTGGGTGCG	GGAGCGATGG	GDCAGGGTGC	1020
GCAATCGGCG	GGCTCCACCA	GGCCGGCTCT	GCTCGCGCGG	GCACCGCTCG	CGCAGGAGCG	1080
TGAAGAGAC	GACGAGGACG	ACTGGGACCA	AGAGCAGCAC	TGGTGAGCTC	CCGTAATGAC	1140
AACAGACTTC	CCGGCCACCC	GGGCGGGAAG	ACTTGCCAAC	ATTTTGCCCA	GGAAGGTAAA	1200
GAGAGGAAGT	AGTCAGCAT	GGCAGAGATG	AAGACCGATG	CCGCTACCCCT	CGCCGAGCAG	1260
GCAGGTAATT	TGAGCGGAT	CTCCGGCGAC	CTGAAABCCC	AGATCGACCA	GSTGGAGTCC	1320
ACGGCAGGTT	CGTTGCAGGG	CCAGTGCGCC	GGCGCGGCGG	GGACGGCGCG	CCAGGCCGCG	1380
GTGGTGCGCT	TCCAAGAAGC	AGCCAATAAG	CAGAAGCAGG	AACTCGACGA	GATCTCAGCG	1440
AATATTCGTC	AGCGCGGCGT	CCATCTACTCG	AGGGCGGAGG	AGGACGAGCA	GCAGGCGCTG	1500
TCTTCGAAA	TGGGCTTCTG	ACCCCTAAT	ACGAAABGAA	ACGAGCAAA	AAGATGACAG	1560
AGCAGCACTG	GAATTTCGCG	GATATCGAGG	CCGCGGCAAG	CGCAATCCAG	GGAAAT	1616

(2) INFORMATION FOR SEQ ID NO:108:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 432 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:108:

CTASTGGATG	GGACCATGGC	CATTTTCGTC	AGTCTCACTG	CCTTCTGTGT	TGACATTTTG	60
GCACCCCGGC	GGAAACGAAG	CACTGGGGTC	GAAGAACGGC	TGGGCTGCCA	TATGTCGGG	120
AGCTTCATA	CCTTCGTGGC	GCCGGAAGAG	CTTGTCTAG	TGGGCCGCCA	TGACAACCTC	180
TCASAGTGCC	CTCAAACGTA	TAAACAGGAS	AAAGGCGGAS	ACCGACGCAA	GGTCGAACTC	240
GCCCCATCCC	GTCTTTCGCT	ATTCTACGCG	AACTCGCGCT	TGCCCCATGC	GAACATCCCA	300
GTGACCTTGC	CTCGGTCGA	AGCCATTGCC	TGACCGGCTT	CGCTGATCGT	CCGCGCCAGG	360
TTCTGCAGCG	CGTTGTTCAg	CTCGGTAGCC	GTGGCTTCCC	ATTTTTCCTC	GACACCTTGG	420
TACGCCTCCG	AA					432

(2) INFORMATION FOR SEQ ID NO:109:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 368 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:109:

Met	Leu	Trp	His	Ala	Met	Pro	Pro	Glu	Xaa	Asn	Thr	Ala	Arg	Leu	Met
1				5					10					15	
Ala	Gly	Ala	Gly	Pro	Ala	Pro	Met	Leu	Ala	Ala	Ala	Ala	Gly	Trp	Gln
			20					25					30		
Thr	Leu	Ser	Ala	Ala	Leu	Asp	Ala	Gln	Ala	Val	Glu	Leu	Thr	Ala	Arg
		35					40				45				
Leu	Asn	Ser	Leu	Gly	Glu	Ala	Trp	Thr	Gly	Gly	Gly	Ser	Asp	Lys	Ala
	50					55					60				
Leu	Ala	Ala	Ala	Thr	Pro	Met	Val	Val	Trp	Leu	Gln	Thr	Ala	Ser	Thr
65					70				75					80	
Gln	Ala	Lys	Thr	Arg	Ala	Met	Gln	Ala	Thr	Ala	Gln	Ala	Ala	Ala	Tyr
				85				90						95	



144

Thr	Gln	Ala	Met	Ala	Thr	Thr	Pro	Ser	Leu	Pro	Glu	Ile	Ala	Ala	Asn		
			100					105					110				
His	Ile	Thr	Gln	Ala	Val	Leu	Thr	Ala	Thr	Asn	Phe	Phe	Gly	Ile	Asn		
		115					120					125					
Thr	Ile	Pro	Ile	Ala	Leu	Thr	Glu	Met	Asp	Tyr	Phe	Ile	Arg	Met	Trp		
		130				135					140						
Asn	Gln	Ala	Ala	Leu	Ala	Met	Glu	Val	Tyr	Gln	Ala	Glu	Thr	Ala	Val		
145				150						155					160		
Asn	Thr	Leu	Phe	Glu	Lys	Leu	Glu	Pro	Met	Ala	Ser	Ile	Leu	Asp	Pro		
			165						170					175			
Gly	Ala	Ser	Gln	Ser	Thr	Thr	Asn	Pro	Ile	Phe	Gly	Met	Pro	Ser	Pro		
			180					185					190				
Gly	Ser	Ser	Thr	Pro	Val	Gly	Gln	Leu	Pro	Pro	Ala	Ala	Thr	Gln	Thr		
		195					200					205					
Leu	Gly	Gln	Leu	Gly	Glu	Met	Ser	Gly	Pro	Met	Gln	Gln	Leu	Thr	Gln		
210					215						220						
Pro	Leu	Gln	Glu	Val	Thr	Ser	Leu	Phe	Ser	Gln	Val	Gly	Gly	Thr	Gly		
225				230						235					240		
Gly	Gly	Asn	Pro	Ala	Asp	Glu	Gln	Ala	Ala	Gln	Met	Gly	Leu	Leu	Gly		
			245					250						255			
Thr	Ser	Pro	Leu	Ser	Asn	His	Pro	Leu	Ala	Gly	Gly	Ser	Gly	Pro	Ser		
			260					265					270				
Ala	Gly	Ala	Gly	Leu	Leu	Arg	Ala	Glu	Ser	Leu	Pro	Gly	Ala	Gly	Gly		
		275					280					285					
Ser	Leu	Thr	Arg	Thr	Pro	Leu	Met	Ser	Gln	Leu	Ile	Glu	Lys	Pro	Val		
290					295						300						
Ala	Pro	Ser	Val	Met	Pro	Ala	Ala	Ala	Ala	Gly	Ser	Ser	Ala	Thr	Gly		
305				310						315					320		
Gly	Ala	Ala	Pro	Val	Gly	Ala	Gly	Ala	Met	Gly	Gln	Gly	Ala	Gln	Ser		
			325					330						335			
Gly	Gly	Ser	Thr	Arg	Pro	Gly	Leu	Val	Ala	Pro	Ala	Pro	Leu	Ala	Gln		
			340				345						350				
Glu	Arg	Gln	Glu	Asp	Asp	Glu	Asp	Asp	Trp	Asp	Glu	Glu	Asp	Asp	Trp		
		355				360						365					

(2) INFORMATION FOR SEQ ID NO:110:

145

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 100 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

```

Met Ala Glu Met Lys Thr Asp Ala Ala Thr Leu Ala Glu Glu Ala Gly
1           5           10           15

Asn Phe Glu Arg Ile Ser Gly Asp Leu Lys Thr Gln Ile Asp Gln Val
          20           25           30

Glu Ser Thr Ala Gly Ser Leu Gln Gly Gln Trp Arg Gly Ala Ala Gly
          35           40           45

Thr Ala Ala Glu Ala Ala Val Val Arg Phe Gln Glu Ala Ala Asn Lys
          50           55           60

Gln Lys Glu Glu Leu Asp Glu Ile Ser Thr Asn Ile Arg Gln Ala Gly
          65           70           75           80

Val Gln Tyr Ser Arg Ala Asp Glu Glu Gln Gln Gln Ala Leu Ser Ser
          85           90           95

Gln Met Gly Phe
          100

```

## (2) INFORMATION FOR SEQ ID NO:111:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 396 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

```

GATCTCCGGC GACCTGAAA CCCAGTCGA CCAGGTGGAG TCGACGGCAG GTTCGTTGCA      60
GGCCAGTGG CCGGGCGGG CGGGACGGC CCGCAGGCC GCGTGGTGC GCTTCCAAAG      120
AGCAGCCAT AGCAGAGGC AGGAATCGA CCAGATCTCG ACGATATTC GTCAGGCCCG      180
CGTCCAATAC TCGAGGGCGG ACGAGGAGCA GCAGCAGCG CTGTCCTCGC AAATGGGCTT      240

```

146

CTGACCCCGCT AATACGAAAA GAACCGGAGC AAAACATCA CAGAGCAGCA GTGGATTTC 300  
 GCGGATATCG AGGCGCGCGC AGCGCAATC CAGGAAATG TCACCTCAT TCATTCCCTC 360  
 CTTGACGAGG GGAACCAATC CCTGACCAAG CTCGCA 396

## (2) INFORMATION FOR SEQ ID NO:112:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 80 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:112:

Ile	Ser	Gly	Asp	Leu	Lys	Thr	Gln	Ile	Asp	Gln	Val	Glu	Ser	Thr	Ala
1				5					10					15	
Gly	Ser	Leu	Gln	Gly	Gln	Trp	Arg	Gly	Ala	Ala	Gly	Thr	Ala	Ala	Gln
			20					25					30		
Ala	Ala	Val	Val	Arg	Phe	Gln	Gln	Ala	Ala	Asn	Lys	Gln	Lys	Gln	Glu
			35				40					45			
Leu	Asp	Glu	Ile	Ser	Thr	Asn	Ile	Arg	Gln	Ala	Gly	Val	Gln	Tyr	Ser
	50					55					60				
Arg	Ala	Asp	Glu	Glu	Gln	Gln	Gln	Ala	Leu	Ser	Ser	Gln	Met	Gly	Phe
65					70				75					80	

## (2) INFORMATION FOR SEQ ID NO:113:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 387 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:113:

GTGGATCCCG ATCCCGTGTT TCGCTATTCT ACGCGAATC GCGCTTCCCC TATGCCAACA 60  
 TCCAGTGAC GTTGCCCTCG GTCGAGCCA TTGCTGACC GCGTTCGCTG ATCGTCCGCG 120  
 CCAGCTTCTG CAGCCCGTTG TTCAGCTCGG TAGCCGTGCG CTCCCATTTT TCGTGACAC 180

```

CCTGGTACGC CTCCGAACCG CTACCGCCCC AGGCCGCTCC GAGCTTGSTC AGGACTGCT      240
TCCCCTCGTC AAGGAGGGGA TGAATGGACG TGACATTCC CTGCATTGCC CTGCGCGCG      300
CCTCGATACC CCGGAAATTC CACTGCTGCT CTGTCATGTT TTTGCTCGST TTCTTTTCGT      360
ATTAGCGGGT CAGAAGCCCA TTTCGCA      387

```

## (2) INFORMATION FOR SEQ ID NO:114:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 272 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

## (1) SEQUENCE DESCRIPTION: SEQ ID NO:114:

```

CGSCACGAGG ATCTGGGTTG GCGCAACGGC GGTGGCGAGG GCTCCGTTC GGGGGCGAGC      60
TGGCGCGCGG ATGCTTCCTC TGCCCGCAGC CCGGCTGGA TGGATCGACC AGTTGCTACC      120
TTCCCGACCT TTCGTTGGGT GTGTGTGCGA TAGCGGTGAC CCGGCGCGCG ACCTCGGGAG      180
TGTTGGGGGG CAGGCGGGGT CGGTGGTTCC GCGCGCGAGC CAGACGCTCT GGACGGAACG      240
GGCGGGGGGT CCGGATTGG CATCTTTGCC CA      272

```

## (2) INFORMATION FOR SEQ ID NO:115:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

## (1) SEQUENCE DESCRIPTION: SEQ ID NO:115:

```

Asp Pro Val Asp Ala Val Ile Asn Thr Thr Cys Asn Tyr Gly Gln Val
 1           5           10           15

Val Ala Ala Leu
          20

```

## (2) INFORMATION FOR SEQ ID NO:116:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

Ala	Val	Glu	Ser	Gly	Met	Leu	Ala	Leu	Gly	Thr	Pro	Ala	Pro	Ser
1				5					10					15

(2) INFORMATION FOR SEQ ID NO:117:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 19 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

Ala	Ala	Met	Lys	Pro	Arg	Thr	Gly	Asp	Gly	Pro	Leu	Glu	Ala	Ala	Lys
1			5					10						15	

Glu Gly Arg

(2) INFORMATION FOR SEQ ID NO:118:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

Tyr	Tyr	Trp	Cys	Pro	Gly	Glu	Pro	Phe	Asp	Pro	Ala	Trp	Gly	Pro
1			5					10					15	

(2) INFORMATION FOR SEQ ID NO:119:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

Asp Ile Gly Ser Gln Ser Thr Glu Asp Gln Gln Asa Ala Val  
1                   5                   10

(2) INFORMATION FOR SEC ID NO: 120:

```
(4) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 13 amino acids
      (B) TYPE: amino acid
      (C) STRANDEDNESS:
      (D) TOPOLOGY: linear
```

(X) SEQUENCE DESCRIPTION; SEQ ID NO:120+

Ala Glu Glu Ser Ile Ser Thr Xaa Glu Xaa Ile Val Pro  
1               5               10

(2) INFORMATION FOR SEQ ID NO:121:

(1.) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 17 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

[81] SEQUENCE DESCRIPTION: SEQ ID NO:121:

Asp Pro Glu Pro Ala Pro Pro Val Pro Thr Thr Ala Ala Ser Pro Pro  
1 5 10 15  
Ser

(2) INFORMATION FOR SEQ ID NO:122:

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 amino acids

150

(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

Ala	Pro	Lys	Thr	Tyr	Xaa	Glu	Gln	Leu	Lys	Gly	Thr	Asp	Thr	Gly
1				5					10					15

(2) INFORMATION FOR SEQ ID NO:123:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 30 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

Asp	Pro	Ala	Ser	Ala	Pro	Asp	Val	Pro	Thr	Ala	Ala	Gln	Leu	Thr	Ser
1				5					10						15
Leu	Leu	Asn	Ser	Leu	Ala	Asp	Pro	Asn	Val	Ser	Phe	Ala	Asn		
				20				25					30		

(2) INFORMATION FOR SEQ ID NO:124:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:

Asp	Pro	Pro	Asp	Pro	His	Gln	Xaa	Asp	Met	Thr	Lys	Gly	Tyr	Tyr	Pro
1				5					10					15	
Gly	Gly	Arg	Arg	Xaa	Phe										
				20											

(2) INFORMATION FOR SEQ ID NO:125:

151

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 7 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:

Asp Pro Gly Tyr Thr Pro Gly  
1 5

(2) INFORMATION FOR SEQ ID NO:126:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ix) FEATURE:

(D) OTHER INFORMATION: /note= "The Second Residue Can Be Either a Pro or Thr"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

Xaa Xaa Gly Phe Thr Gly Pro Gln Phe Tyr  
1 5 10

(2) INFORMATION FOR SEQ ID NO:127:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 9 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ix) FEATURE:

(D) OTHER INFORMATION: /note= "The Third Residue Can Be Either a Gln or Leu"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:

Xaa Pro Xaa Val Thr Ala Tyr Ala Gly  
1 5



## (2) INFORMATION FOR SEQ ID NO:128:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 9 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

Xaa Xaa Xaa Glu Lys Pro Phe Leu Arg  
1 5

## (2) INFORMATION FOR SEQ ID NO:129:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

Xaa Asp Ser Glu Lys Ser Ala Thr Ile Lys Val Thr Asp Ala Ser  
1 5 10 15

## (2) INFORMATION FOR SEQ ID NO:130:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

Ala Gly Asp Thr Xaa Ile Tyr Ile Val Gly Asn Leu Thr Ala Asp  
1 5 10 15

## (2) INFORMATION FOR SEQ ID NO:131:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 amino acids

153

(E) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

Ala Pro Glu Ser Gly Ala Gly Leu Gly Gly Thr Val Gln Ala Gly  
 1                      5                      10                      15

(2) INFORMATION FOR SEQ ID NO:132:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 21 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

Xaa Tyr Ile Ala Tyr Xaa Thr Thr Ala Gly Ile Val Pro Gly Lys Ile  
 1                      5                      10                      15

Asn Val His Leu Val  
 20

(2) INFORMATION FOR SEQ ID NO:133:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 982 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

GCAACCGCTGT CGTGGCCTTT GCGGTGATCG GTTGGCCTC GCTGGCGGTG GCGGTGGCGG	60
TCACCATCCG ACCGACCGCG GCCTCAAAAC CGGTAGAGGG ACACCAAAAC GCCAGGCCAG	120
GGAAGTTCTT GCCCTTGTTG CCGACGCAAC AGCAGCGCGC GGTCCCGCCG CCTCCGCCCG	180
ATGATCCCAAC CGCTGGATTC CAGGGCGGCA CCATTCGGGC TGTACAGAAC GTGGTCCCGC	240

```

GGCCCGGTAC CTCACCCGGG GTGGGTGGGA CGCCGCTTC GCCTGCGCCG GAAGCGCCGG 300
CCGTGCCCCG TGTTGTGCTT GCCCCGGTGC CAATCCCGGT CCGGATCATC ATTCCCCCGT 360
TCCCGGGTTG GCAGCCTGGA ATGCGGACCA TCCCCACCGC ACCGCGGAGG ACGCCGGTGA 420
CCACGTCGSC GACGACGCCG CCGACACCGC CGCGACCCAC GCCGCTGACC ACGCCGCCAA 480
CGACGCGGCC GACCACGCCG GTGACACCGC CGCCACCGAC GCTGCGGACC ACGCCGGTGA 540
CCACGCCACC ACGACCGCTC GCCCCGACGA CGCTGCGCCG GACGACCGTC GCTCCGACCA 600
CGCTGCGCCG GACCGCGCTC GCTCCGACCA CGCCGACCGC GACGACCGTC GCTCCGACCA 660
CGACCGACCA GCGGACCGAA CAACCAACCG AACAGATGCC AACCCAGCAG CAGACCGTGG 720
CCGCGCAGAC GGTGGCGCCG GCTCCGACCG CGCCGTCGCG TGGCGCGCAC GGCAGCGCGG 780
GGGCGGACTT ATTCGGCGGG TCTGATCAC GGTGCGCGCT TCACTACGGT CGGAGGACAT 840
GGCGGTCAT GCGGTGACGG TGGTCTGCC CTGTCTCACC GA 882

```

## (2) INFORMATION FOR SEQ ID NO:134:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 815 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:

```

CCATCAACCA ACCGCTGCGC CCGCCGCGGC CCGCGGATCC GCGTCCGCG CCACGCCCGC 60
CGGTGCTTCC GGTGCCCCCG TTGCGCGCGT CGCCGCGGTC GCGGCGGACC GGTGGGTGC 120
CTAGGGCGCT GTTACCGCCC TGGTTGCGCG GAGCGCGCGC GGCACCGCG GTACCGCGCA 180
TGGCGCGGTT GCGCGCGCGC GCACCGTTGC CACCGTTGCC ACCGTTGCUA CGCTTGCGCA 240
CCAGCCACCC GCGCGGACCA CCGGACCGGC CGCGCGCGCC CGCACCGCUG GCGTGCCTGT 300
TCGTGCGCGT ACCGCGCGCA CCGCGGTTGC CGCGCTACC GCGGACCGAA CTACCGCGCG 360
ACGCGCGCGT CCGCGCGCGC CCGCGCGCAC CGCGATTGCC ACCGCGGTC CCGCGCGCGT 420
GGAGTGCGCC GATTAGGCGA CTGACCGCGC CAACCAACCG AAGTACTCTC GGTACCGGAG 480
CACTTCGCGA CGACACCGA GCACCGCGTT GTCGCGCGAC TGGGTGNAAT GCGAGCGCAT 540

```

155

```

AGCGGCTAGC TGTGGGCTEC GTTCAAGCTC GATCATGATC TCGAGGTGAC CSTGACCGCG 600
CCCCCGGAAG GAGGGGCTCA ACTCGGGCTT GAGGCGATCG GCGATCGTTT GGGGCAGTGC 660
CCAGGCCAAT ACGGGGATAC CGSSTGTGNA AGCCGCGCGG AGCGCAGCTT CGGTTGCGCG 720
ACNGTGGTCS GCGTGGGCTG TTACGCCGTT GTGNTGGAAC ACGAGTAGCA GGTCTGCTCC 780
GGCGAGGGCA TCCACCACGC GTTGCGTCAG CTCGT 815

```

(2) INFORMATION FOR SEQ ID NO:135:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1152 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:

```

ACCAGCCGCC GCGTGAGGTC TCAGATCAGA GAGTCTCCGG ACTCACCBSG GCGGTTCAAG 60
CTTCTCCAG AACAACTGCT GAAGATCCTC GCGCGGAAA CAGGCGCTGA TTGACGCTC 120
TATGACCGGT TGAACGACGA GATCATCCGG CAGATTGATA TGGCACGGCT GGGCTAACAG 180
GTGCGCAAGA TCGTGCACT GTATGTCTCG GACTCCCTGT CCGGGATCAG CTTTCCCGAC 240
GGCGGGGTGA TCGTGTGGAG CGAGGAGCTC GCGCAGAGGC AGTATCCGAT CGAGACGCTG 300
GACGGCATCA CCTGTTTGG GCGCGGAGG ATGACAACGC CTTTCATCTT TGAGATGCTC 360
AAGCGTAGC GCGACATCCA GTCTTTCAG ACCGACGGC ACTACLAGGG CCGGATCTCA 420
ACACCCGACG TGTATACGC GCGCGGCTC CGTCAACAG TTCACCGCAC CGACGATCTT 480
GCGTTCTGCC TGTGTTAAG CAGCGGATC GTGTGAGGA AGATCCTGAA TCAGCAGGCC 540
TTGATTCGGG CACACACCTC GGGGCAAGAC GTTGCTGAGA GCATCCGCAC GATCAAGCAC 600
TGGCTGGCT GGGTCGATCG ATCGGGCTCC CTGGCGAGT TGAACGGGTT CGAGCGAAT 660
GCGCAAAAG CTACTTCAC GCGCTGGGG CATCTCTCC CCGAGGATT CGCATTCACG 720
GCGGCTCGA CTCGGGCGCC GTTGAGCGCC TTCAACTGAA TGGTCAGCTT CGGCTATTCT 780
CTGCTGTACA AGAATATCAT AGGGGCGATC GAGCGTCACA GCGTGAACGC GTATATCGGT 840
TTCTACACC AGGATTCAG AGGGCAGCA ACCTCTCGTG CCGAATTCGG CACGAGCTCC 900

```

156

```

GCTGAACCCG CTGCCCCGCT GCTCAGTGCC CTTACGTAAT CCGCTGCGCC CAGGCCGGCC      960
CGCCGGGCGA ATACGAGCAG ATCGGACAGC GAATTGCGGC CCAGCCGCTT GGAGCCGTGC      1020
ATACCGCCCG CACACTCACC GGCAGCGAAC AGGCTGGCA CCGTGGCGC GCCGGTGTCC      1080
GGCTCTACTT CGACACCGCC CATCAGCTAG TGACAGCTCG GCGCGACTTC CATTCGCTCC      1140
GTTCCGCACG AG                                     1152

```

(2) INFORMATION FOR SEQ ID NO:136:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 655 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:136:

```

CTCTGCCCCG TTCGGCAGGG TGTACTTGCC GGTGGTGTAN GCGGATGAG TCCCGACGAC      60
CAGCAATGCG GCAACAGCAC GATCCCGCTT CAGGACGGCC ACCCGGTCCA CTTGGGCGAT      120
CCGCTCGACT CCGCCCTGGG CCGCTCTTTC CTTGGGCAGG CTCATCCGAC GTGTTTCCGC      180
CTTGCTTTTC CGCCATTATG CCGGCGCGCC GCTTCGGCG GCGGTATGG CCGAAGCTCG      240
ATCAGCAGAC CCGAGATACG GGTCTGTGCA AGCTTTTIGA GGTTCGCGCG GGGCAGCTTC      300
GCCGCGATT CTACTAGCGA GAAGTCTGGC CCGATACCGA TCTGACCGAA GTCGCTGCGG      360
TGCAGCCGAC CTCATTGGC GATGGCGCGG ACGATGCGCG CTGGACCGAT CTTGTGCGCG      420
TTGCGGACCG CGACCGGTA GGTGGTCAAG TCCGCTCTAC GCTTGGGCGT TTGCGGACCG      480
TCCGAGCCT GGTGCGGCTT GCGCGCGAA AGCGCGGGT CCGGTGCCAT CAGGAATGCC      540
TCACCGCCCC GCACTGCAC GCGCACTGCC CCGCGATGT CAGCCATCG GACATCATGC      600
TCGCTTCAT ACTCTCGAC CACTCGCGG AACACTCGA TTCCCGGACC GCGCA      655

```

(2) INFORMATION FOR SEQ ID NO:137:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 267 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:

Asn	Ala	Val	Val	Ala	Phe	Ala	Val	Ile	Gly	Phe	Ala	Ser	Leu	Ala	Val	1	5	10	15
Ala	Val	Ala	Val	Thr	Ile	Arg	Pro	Thr	Ala	Ala	Ser	Lys	Pro	Val	Glu	20	25	30	
Gly	His	Gln	Asn	Ala	Gln	Pro	Gly	Lys	Phe	Met	Pro	Leu	Leu	Pro	Thr	35	40	45	
Gln	Gln	Gln	Ala	Pro	Val	Pro	Pro	Pro	Pro	Pro	Asp	Asp	Pro	Thr	Ala	50	55	60	
Gly	Phe	Gln	Gly	Gly	Thr	Ile	Pro	Ala	Val	Gln	Asn	Val	Val	Pro	Arg	65	70	75	80
Pro	Gly	Thr	Ser	Pro	Gly	Val	Gly	Gly	Thr	Pro	Ala	Ser	Pro	Ala	Pro	85	90	95	
Glu	Ala	Pro	Ala	Val	Pro	Gly	Val	Val	Pro	Ala	Pro	Val	Pro	Ile	Pro	100	105	110	
Val	Pro	Ile	Ile	Ile	Pro	Pro	Phe	Pro	Gly	Trp	Gln	Pro	Gly	Met	Pro	115	120	125	
Thr	Ile	Pro	Thr	Ala	Pro	Pro	Thr	Thr	Pro	Val	Thr	Thr	Ser	Ala	Thr	130	135	140	
Thr	Pro	Pro	Thr	Thr	Pro	Pro	Thr	Thr	Pro	Val	Thr	Thr	Pro	Pro	Thr	145	150	155	160
Thr	Pro	Pro	Thr	Thr	Pro	Val	Thr	Thr	Pro	Pro	Thr	Thr	Pro	Pro	Thr	165	170	175	
Thr	Pro	Val	Thr	Thr	Pro	Pro	Thr	Thr	Val	Ala	Pro	Thr	Thr	Val	Ala	180	185	190	
Pro	Thr	Thr	Val	Ala	Pro	Thr	Thr	Val	Ala	Pro	Thr	Thr	Val	Ala	Pro	195	200	205	
Ala	Thr	Ala	Thr	Pro	Thr	Thr	Val	Ala	Pro	Gln	Pro	Thr	Gln	Gln	Pro	210	215	220	
Thr	Gln	Gln	Pro	Thr	Gln	Gln	Met	Pro	Thr	Gln	Gln	Gln	Thr	Val	Ala	225	230	235	240
Pro	Gln	Thr	Val	Ala	Pro	Ala	Pro	Gln	Pro	Pro	Ser	Gly	Gly	Arg	Asn	245	250	255	

158

Gly Ser Gly Gly Gly Asp Leu Phe Gly Gly Phe  
 260 265

## (2) INFORMATION FOR SEQ ID NO:138:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 174 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:138:

```

Ile Asn Gln Pro Leu Ala Pro Pro Ala Pro Pro Asp Pro Pro Ser Pro
1           5           10           15

Pro Arg Pro Pro Val Pro Pro Val Pro Pro Leu Pro Pro Ser Pro Pro
20           25           30

Ser Pro Pro Thr Gly Trp Val Pro Arg Ala Leu Leu Pro Pro Trp Leu
35           40           45

Ala Gly Thr Pro Pro Ala Pro Pro Val Pro Pro Met Ala Pro Leu Pro
50           55           60

Pro Ala Ala Pro Leu Pro Pro Leu Pro Pro Leu Pro Pro Leu Pro Thr
65           70           75           80

Ser His Pro Pro Arg Pro Pro Ala Pro Pro Ala Pro Pro Ala Pro Pro
85           90           95

Ala Cys Pro Phe Val Pro Val Pro Pro Ala Pro Pro Leu Pro Pro Ser
100          105          110

Pro Pro Thr Glu Leu Pro Ala Asp Ala Ala Cys Pro Pro Ala Pro Pro
115          120          125

Ala Pro Pro Leu Ala Pro Pro Ser Pro Pro Ala Gly Ser Ala Ala His
130          135          140

Arg Ala Leu Thr Gly Ala Thr Ser Ala Ser Thr Leu Gly His Arg Ala
145          150          155          160

Leu Pro Asp Asp Thr Thr Ala Arg Gly Cys Arg Arg Thr Gly
165          170

```

## (2) INFORMATION FOR SEQ ID NO:139:

## (i) SEQUENCE CHARACTERISTICS:





100

## (2) INFORMATION FOR SEQ ID NO:141:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 53 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PCR primer"

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Mycobacterium tuberculosis

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

GGATCATAT GGGCATCAT CATCATCATC ACCTGATCGA CATCATCGCG ACC

53

## (2) INFORMATION FOR SEQ ID NO:142:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PCR Primer"

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Mycobacterium tuberculosis

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:142:

CCTGAATTCA GGCTCGGTT GCGCCGGCCT CATCTTGAAC GA

42

## (2) INFORMATION FOR SEQ ID NO:143:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PCR Primer"

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

GGATCCTGCA GGCTCGAAAC CACCGAGCGG T

31

(2) INFORMATION FOR SEQ ID NO:144:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc ~ "PCR primer"

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

CTCTGAATTC AGCGCTGGAA ATCGTCGCGA T

31

(2) INFORMATION FOR SEQ ID NO:145:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc ~ "PCR primer"

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

GGATCCAGCG CTCGATGAA GACCGATGCC GCT

33

(2) INFORMATION FOR SEQ ID NO:146:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

{1i} MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "PCR primer"

{vi} ORIGINAL SOURCE:  
(A) ORGANISM: Mycobacterium tuberculosis

{xi} SEQUENCE DESCRIPTION: SEQ ID NO:146:

GAGAGAATTC TCAGAGGCCC ATTTGCGAGG ACA

33

{2} INFORMATION FOR SEQ ID NO:147:

{1} SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1993 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

{1i} MOLECULE TYPE: DNA (genomic)

{vi} ORIGINAL SOURCE:  
(A) ORGANISM: Mycobacterium tuberculosis

{lx} FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 152..1273

{xi} SEQUENCE DESCRIPTION: SEQ ID NO:147:

TGTTCTTCGA CCGCAGGCTG GTGAGGAAAG GGGCCACCGA ACAGCTGTTT TCCTGCGCGA	60
AGCATGCGGA AAGCGCCCGA TACGTGCGCG GACTGTGCGG GGACCTCAAG GACGCGAAGC	120
CGCGAATTG AAGAGCACAG AAGGTATGG C GTG AAA ATT CGT TTG CAT ACG	172
Val Lys Ile Arg Leu His Thr	
1 5	
CTG TTG GCG GTG TTG ACC GGT GCG CCG CTG CTG CTA GCA GCG GCG GGC	220
Leu Leu Ala Val Leu Thr Ala Ala Pro Leu Leu Leu Ala Ala Ala Gly	
10 15 20	
TGT GCG TCG AAA CCA CCG AGC GGT TCG CGT GAA ACG GCG GCG GCG GCG	268
Cys Gly Ser Lys Pro Pro Ser Gly Ser Pro Glu Thr Gly Ala Gly Ala	
25 30 35	
GGT ACT GTC GCG ACT ACC CCG GCG TCG TCG CCG GTG ACG TTG GCG GAG	316
Gly Thr Val Ala Thr Thr Pro Ala Ser Ser Pro Val Thr Leu Ala Glu	
40 45 50 55	
ACC GGT ACC ACG CTG CTC TAC CCG CTG TTC AAC CTG TCG GGT CCG GCG	364
Thr Gly Ser Thr Leu Leu Tyr Pro Leu Phe Asn Leu Trp Gly Pro Ala	
60 65 70	

TTF CAC CAG AGG TAT CCG AAC GTC ACG ATC ACC GCT CAG GGC ACC GGT	412
Phe His Glu Arg Tyr Pro Asn Val Thr Ile Thr Ala Gln Gly Thr Gly	
75 80 85	
TCT GGT GCC GGG ATC GCG CAG GCC GCC GCC GGG ACG GTC AAC ATT GGG	460
Ser Gly Ala Gly Ile Ala Gln Ala Ala Ala Gly Thr Val Asn Ile Gly	
90 95 100	
GCC TCC CAC GCC TAT CTC TCG GAA GGT GAT ATC GCC GCG CAC AAG GGG	508
Ala Ser Asp Ala Tyr Leu Ser Glu Gly Asp Met Ala Ala His Lys Gly	
105 110 115	
CTG ATG AAC ATC GCG CTA GCC ATC TCC GCT CAG CAG GTC AAC TAC AAC	556
Leu Met Asn Ile Ala Leu Ala Ile Ser Ala Gln Gln Val Asn Tyr Asn	
120 125 130 135	
CTG CCC GGA GTG AGC GAG CAC CTC AAG CTG AAC CGA AAA GTC CTG GCG	604
Leu Pro Gly Val Ser Glu His Leu Lys Leu Asn Gly Lys Val Leu Ala	
140 145 150	
GCC ATG TAC CAG GGC ACC ATC AAG ACC TGG GAC GAC CCG CAG ATC GCT	652
Ala Met Tyr Gln Gly Thr Ile Lys Thr Trp Asp Asp Pro Gln Ile Ala	
155 160 165	
GCG CTC AAC CCC GGC GTG AAC CTG CCC GGC ACC GCG GYA GTT CCG CTG	700
Ala Leu Asn Pro Gly Val Asn Leu Pro Gly Thr Ala Val Val Pro Leu	
170 175 180	
CAC CCG TCC GAC GGG TCC GGT GAC ACC TTC TTG TTC ACC CAG TAC CTG	748
His Arg Ser Asp Gly Ser Gly Asp Thr Phe Leu Phe Thr Gln Tyr Leu	
185 190 195	
TCC AAG CAA GAT CCC GAG GGC TGG GGC AAG TCG CCC GGC TTC GGC ACC	796
Ser Lys Gln Asp Pro Glu Gly Trp Gly Lys Ser Pro Gly Phe Gly Thr	
200 205 210 215	
ACC GTC GAC TTC CCG GCG GTG CCG GGT CCG CTG GGT GAG AAC GGC AAC	844
Thr Val Asp Phe Pro Ala Val Pro Gly Ala Leu Gly Glu Asn Gly Asn	
220 225 230	
GCC GGC ATG GTG ACC GGT TGC GCC GAG ACA CCG GGC TGC CTG CCC TAT	892
Gly Gly Met Val Thr Gly Cys Ala Glu Thr Pro Gly Cys Val Ala Tyr	
235 240 245	
ATC GCC ATC AGC TTC CTC GAC CAG CCC AGT CAA CCG GGA CTC GGC GAG	940
Ile Gly Ile Ser Phe Leu Asp Gln Ala Ser Gln Arg Gly Leu Gly Glu	
250 255 260	
GCC CAA CTA GCC AAT AGC TCT GGC AAT TTC TTG TTG CCC GAC GCG CAA	988
Ala Gln Leu Gly Asn Ser Ser Gly Asn Phe Leu Leu Pro Asp Ala Gln	
265 270 275	
AGC ATT CAG GCC GCG GCG GCT GGC TTC GCA TCG AAA ACC CCG GCG AAC	1036
Ser Ile Gln Ala Ala Ala Ala Gly Phe Ala Ser Lys Thr Pro Ala Asn	

280	285	290	295	
CAG GCG ATT TCG ATG ATC SAC GGG CCC GCG CCG GAC GGC TAC CCG ATC	Gln Ala Ile Ser Met Ile Asp Gly Pro Ala Pro Asp Gly Tyr Pro Ile	1684		
	300	305	310	
ATC AAC TAC GAG TAC GCC ATC CTC AAC AAC CCG CAA AAG GAC GCC GCC	Ile Asn Tyr Gln Tyr Ala Ile Val Asn Asn Arg Gln Lys Asp Ala Ala	1132		
	315	320	325	
ACC GCG CAG ACC TTG CAG GCA TTT CTG CAC TGG GCG ATC ACC GAC GGC	Thr Ala Gln Thr Leu Gln Ala Phe Leu His Trp Ala Ile Thr Asp Gly	1180		
	330	335	340	
AAC AAG GCC TCG TTC CTC GAC CAG GTT CAG TTC CAG CCG CTG CCG CCC	Asn Lys Ala Ser Phe Leu Asp Gln Val His Phe Gln Pro Leu Pro Pro	1228		
	345	350	355	
GCG GTG GTG AAG TTG TCT GAC CCG TTG ATC GCG ACG ATT TCC AGC	Ala Val Val Lys Leu Ser Asp Ala Leu Ile Ala Thr Ile Ser Ser	1272		
	360	365	370	
TAGCCTCGTT GACGACCAGS CGACAGCAAC CTCGTCGGG CCATCGGCGT GCTTTGCGGA		1333		
GCATGCTGAC CCGTGCCGCT GAAGTCGGCC GCGCTGCGCC GCCCATCCGG TGGTTGCGTG		1393		
GGATAGGTGC GTGATCCCG CTGCTTGCGC TGGTCTTGGT GCTGGTGGTG CTGCTCATCG		1453		
AGCGGATGGG TCGGATCAGG CTCACCGGCT TGCATTTCTT CACCGCCACC GAATGGGAATC		1513		
CAGGCAACAC CTACGGCGAA ACCGTTGTCA CCGACGCTC GCCCATCCGG TGGGCGGCTA		1573		
CTACGGGGCG TTGCGGCTGA TCGTCGGGAC GCTGGCGACC TGGGCAATCG GCGTGATCAT		1633		
CGCGSTGCGS GTCTCTGTAG GAGCGGCGCT GGTGATCGTG GAACGGCTGC CGAAACGGTT		1693		
GGCCGAGGCT GTGGGAATAG TCGTGGAATT GCTCGCCGGA ATCCCGAGCG TGCTCGTCGG		1753		
TTTGTGGGCG GCAATGACGT TCGGGCCGTT CATGCTCAT CACATCGCTC CGGTGATCGC		1813		
TCACAACGCT CCGGATGTGC CGGTGCTGAA CTACTTGCAC GCGGACCCGG GCAACGGGGA		1873		
GGGCATGTTG GTGTCCGCTC TGGTGTGGC GGTGATGGTC GTTCCCATTA TCGCCACGAC		1933		
CACTCATGAC CTGTTCCGCG AGGTGCGGCT GTTCCCGCG GAGGGCGGGA TCGGGAATTC		1993		

(2) INFORMATION FOR SEQ ID NO:148:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 374 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

165

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:

Val	Lys	Ile	Arg	Leu	His	Thr	Leu	Leu	Ala	Val	Leu	Thr	Ala	Ala	Pro	1	5	10	15
Leu	Leu	Leu	Ala	Ala	Ala	Gly	Cys	Gly	Ser	Lys	Pro	Pro	Ser	Gly	Ser	20	25	30	
Pro	Glu	Thr	Gly	Ala	Gly	Ala	Gly	Thr	Val	Ala	Thr	Thr	Pro	Ala	Ser	35	40	45	
Ser	Pro	Val	Thr	Leu	Ala	Glu	Thr	Gly	Ser	Thr	Leu	Leu	Tyr	Pro	Leu	50	55	60	
Phe	Asn	Leu	Trp	Gly	Pro	Ala	Phe	His	Glu	Arg	Tyr	Pro	Asn	Val	Thr	65	70	75	80
Ile	Thr	Ala	Gln	Gly	Thr	Gly	Ser	Gly	Ala	Gly	Ile	Ala	Gln	Ala	Ala	85	90	95	
Ala	Gly	Thr	Val	Asn	Ile	Gly	Ala	Ser	Asp	Ala	Tyr	Leu	Ser	Glu	Gly	100	105	110	
Asp	Met	Ala	Ala	His	Lys	Gly	Leu	Met	Asn	Ile	Ala	Leu	Ala	Ile	Ser	115	120	125	
Ala	Gln	Gln	Val	Asn	Tyr	Asn	Leu	Pro	Gly	Val	Ser	Glu	His	Leu	Lys	130	135	140	
Leu	Asn	Gly	Lys	Val	Leu	Ala	Ala	Met	Tyr	Gln	Gly	Thr	Ile	Lys	Thr	145	150	155	160
Trp	Asp	Asp	Pro	Gln	Ile	Ala	Ala	Leu	Asn	Pro	Gly	Val	Asn	Leu	Pro	165	170	175	
Gly	Thr	Ala	Val	Val	Pro	Leu	His	Arg	Ser	Asp	Gly	Ser	Gly	Asp	Thr	180	185	190	
Phe	Leu	Phe	Thr	Gln	Tyr	Leu	Ser	Lys	Glu	Asp	Pro	Glu	Gly	Trp	Gly	195	200	205	
Lys	Ser	Pro	Gly	Phe	Gly	Thr	Thr	Val	Asp	Phe	Pro	Ala	Val	Pro	Gly	210	215	220	
Ala	Leu	Gly	Glu	Asn	Gly	Asn	Gly	Gly	Met	Val	Thr	Gly	Cys	Ala	Gln	225	230	235	240
Thr	Pro	Gly	Cys	Val	Ala	Tyr	Ile	Gly	Ile	Ser	Phe	Leu	Asp	Gln	Ala	245	250	255	
Ser	Gln	Arg	Gly	Leu	Gly	Glu	Ala	Gln	Leu	Gly	Asn	Ser	Ser	Gly	Asn	260	265	270	

166

Phe Leu Leu Pro Asp Ala Gln Ser Ile Gln Ala Ala Ala Ala Gly Phe  
 275 280 285  
 Ala Ser Lys Thr Pro Ala Asn Gln Ala Ile Ser Met Ile Asp Gly Pro  
 290 295 300  
 Ala Pro Asp Gly Tyr Pro Ile Ile Asn Tyr Gln Tyr Ala Ile Val Asn  
 305 310 315 320  
 Asn Arg Gln Lys Asp Ala Ala Thr Ala Gln Thr Leu Gln Ala Phe Leu  
 325 330 335  
 His Trp Ala Ile Thr Asp Gly Asn Lys Ala Ser Phe Leu Asp Gln Val  
 340 345 350  
 His Phe Gln Pro Leu Pro Pro Ala Val Val Lys Leu Ser Asp Ala Leu  
 355 360 365  
 Ile Ala Thr Ile Ser Ser  
 370

## (2) INFORMATION FOR SEQ ID NO:149:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1993 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:149:

TGTTCTTCGA CCGCAGGCTG GTGAGGAGG GGGCCACCGA ACAGCTGTTC TGCTGCGCGA	60
AGCATGCGGA AACCGCCCGA TACGTGCGCG GACTGTGCGG GGAGGTCAAG GACGCCAAGC	120
GCGGAAATTG AAGAGCACAG AAAGGTATGG CCGGAAAATT CGTTTGATA CGCTGTGGGC	180
CGTGTGACC GGTGCGCGGC TGCTGCTAGC AGCGCGGGGC TGTGGCTCGA AACCACTGAG	240
CGGTCGCTT GAAACGGGGG CCGGCGCGCG TACTGTGCGG ACTACCCCGG CGTGTGCGGC	300
GGTGACGTTG GCGGAGACCG GTAGCAGCTT GCTCTACCGG CTGTTCAACC TGTGGGGTCC	360
GGCTTTTCAC GAGAGGTATC CGAGCTTCAC GATCACCGCT CAGGGCACCG GTTCTGGTGC	420
CGGATCGCG CAGCGCGCGG CCGGAGCGGT CAACATTGGG GGCTCGGACG CCTATCTGTC	480
GGAAGGTGAT ATGGCCGCGC ACAAGGGGCT GATGACATC GCGGTAGCCA TCTCCGCTCA	540
GCAGGTCAAC TACAACCTGC CCGGAGTGGG CGAGCACTTC AAGCTGAACG GAAAACTCTT	600

GGCGGCCATG	TACCAAGGGA	CCATCAAAAC	CTGGGACGAC	CCGCAGATCG	CTGCGCTCAA	660
CCCCGGCCTG	AACCTGCCCC	GCACCGCGGT	AGTTCCGCTG	CACCGCTCCG	ACGGGTCCGG	720
TCACACCTTC	TTGTTCACCC	AGTACCTGTC	CAAGCAAGAT	CCCGAGGECT	GGGGCAAGTC	780
GCCCCGCTTC	GGCACCACCG	TGCACCTCCC	GGCGGTGCCG	GGTGGCTCGG	GTGAGAACCG	840
CAACGGCGGC	ATGGTGACCG	GTTGCGCCCA	GACACCGGGC	TCCGTGGCCT	ATATCGGCAT	900
CAGCTTCCTC	CACCAAGCCA	GTCAACCGCG	ACTCGCGCAG	GCCCAACTAG	GCAATAGCTC	960
TGGCAATTTT	TTGTTGCCCC	ACGCGCAAAG	CAATCAGGCC	CCCGCGGCTG	GCTTCGCAAT	1020
GAAAACCCCG	GCGACCCAGG	CGATTTGCGT	GATCGAGCGG	CCCGCCCCCG	ACGGCTACCC	1080
GATCATCAAC	TACGAGTACG	CCATCTGTAA	CAACCGGCAG	AAGGACGCGG	CCACCGCCCA	1140
GACCTTGCGG	GCATTTCTGC	ACTGGGCGAT	CACCGACGGC	AACAAGGCGT	CGTTCTTCGA	1200
CCAGGTTTCT	TTCCAGCCCG	TGCGCGCCCG	GGTGGTGAAG	TGTCTGACG	CGTTGATCGC	1260
GACGATTTTC	AGCTAGCCCT	GTTGACCACC	ACCGACACGC	AACCTCCGTC	GGGCCATCCG	1320
GCTGCTTTTC	CGACCATGCT	GGCCCGTGGC	GGTGAAGTGG	CCCCCGGCTG	CCCGGCGATC	1380
CGGTGGTTGG	GTGGGATAGG	TGGGTTGATC	CCGCTGCTTG	CGCTGGTCTT	GGTGGTGGTG	1440
GTGCTGGTCA	TCGAGGCGAT	GGTTCGATC	AGGCTCAACG	GGTTGCATTT	CTTCACCCGC	1500
ACCGAATGGA	ATCCAGGCAG	CACCTACGGC	GAAACGGTTG	TCACCGACCG	GTCCGCCATC	1560
CGGTGGGGCG	CTACTACGGG	GCCTGCCCCC	TGATCGTCCG	GACGCTGGCG	ACCTCGGCCA	1620
TCGCCCTCAT	CATCCCGGTG	CCGGTCTCTG	TAGGAGCGGC	GCTGGTGATC	GTGGAACGGC	1680
TGCCGAACCG	GTTCGCCGAG	GCCTGCGGAA	TAGTCTTGGA	ATTGCTCGCC	GGAATCCCCA	1740
GCGTGGTGGT	CGGTTTGTGG	GGGCAATGA	CGTTGCGGCG	GTTCATCGCT	CATCACAATG	1800
CTCCGGTGAT	CGCTCACAAC	GCTCCCGATG	TGCCGGTGCT	GAACACTTTG	CGCGCGGACC	1860
CGGGCAACCG	GGAGGGCAGT	TTGGTGTCCG	GTCTGGTGTG	GGCGGTGATG	GTCTTTCCCA	1920
TTATCGCCAC	CACCACTCAT	GACCTGTTCG	GGCAGGTGCC	GCTGTTCCCG	CGCGAGGGCG	1980
CGATCGGGAA	TTC					1993

(2) INFORMATION FOR SEQ ID NO:156:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 374 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear



(x1) SEQUENCE DESCRIPTION: SEQ ID NO:150:

Met	Lys	Ile	Arg	Leu	His	Thr	Leu	Leu	Ala	Val	Leu	Thr	Ala	Ala	Pro	1	5	10	15
Leu	Leu	Leu	Ala	Ala	Ala	Gly	Cys	Gly	Ser	Lys	Pro	Pro	Ser	Gly	Ser	20	25	30	
Pro	Glu	Thr	Gly	Ala	Gly	Ala	Gly	Thr	Val	Ala	Thr	Thr	Pro	Ala	Ser	35	40	45	
Ser	Pro	Val	Thr	Leu	Ala	Glu	Thr	Gly	Ser	Thr	Leu	Leu	Tyr	Pro	Leu	50	55	60	
Phe	Asn	Leu	Trp	Gly	Pro	Ala	Phe	His	Glu	Arg	Tyr	Pro	Asn	Val	Thr	65	70	75	80
Ile	Thr	Ala	Gln	Gly	Thr	Gly	Ser	Gly	Ala	Gly	Ile	Ala	Gln	Ala	Ala	85	90	95	
Ala	Gly	Thr	Val	Asn	Ile	Gly	Ala	Ser	Asp	Ala	Tyr	Leu	Ser	Glu	Gly	100	105	110	
Asp	Met	Ala	Ala	His	Lys	Gly	Leu	Met	Asn	Ile	Ala	Leu	Ala	Ile	Ser	115	120	125	
Ala	Gln	Gln	Val	Asn	Tyr	Asn	Leu	Pro	Gly	Val	Ser	Glu	His	Leu	Lys	130	135	140	
Leu	Asn	Gly	Lys	Val	Leu	Ala	Ala	Met	Tyr	Gln	Gly	Thr	Ile	Lys	Thr	145	150	155	160
Trp	Asp	Asp	Pro	Gln	Ile	Ala	Ala	Leu	Asn	Pro	Gly	Val	Asn	Leu	Pro	165	170	175	
Gly	Thr	Ala	Val	Val	Pro	Leu	His	Arg	Ser	Asp	Gly	Ser	Gly	Asp	Thr	180	185	190	
Phe	Leu	Phe	Thr	Gln	Tyr	Leu	Ser	Lys	Gln	Asp	Pro	Glu	Gly	Trp	Gly	195	200	205	
Lys	Ser	Pro	Gly	Phe	Gly	Thr	Thr	Val	Asp	Phe	Pro	Ala	Val	Pro	Gly	210	215	220	
Ala	Leu	Gly	Glu	Asn	Gly	Asn	Gly	Gly	Met	Val	Thr	Gly	Cys	Ala	Glu	225	230	235	240
Thr	Pro	Gly	Cys	Val	Ala	Tyr	Ile	Gly	Ile	Ser	Phe	Leu	Asp	Gln	Ala	245	250	255	

Ser Gln Arg Gly Leu Gly Gln Ala Gln Leu Gly Asn Ser Ser Gly Asn  
 260 265 270  
 Phe Leu Leu Pro Asp Ala Gln Ser Ile Gln Ala Ala Ala Gly Phe  
 275 280 285  
 Ala Ser Lys Thr Pro Ala Asn Gln Ala Ile Ser Met Ile Asp Gly Pro  
 290 295 300  
 Ala Pro Asp Gly Tyr Pro Ile Ile Asn Tyr Glu Tyr Ala Ile Val Asn  
 305 310 315 320  
 Asn Arg Gln Lys Asp Ala Ala Thr Ala Gln Thr Leu Gln Ala Phe Leu  
 325 330 335  
 His Trp Ala Ile Thr Asp Gly Asn Lys Ala Ser Phe Leu Asp Gln Val  
 340 345 350  
 His Phe Gln Pro Leu Pro Pro Ala Val Val Lys Leu Ser Asp Ala Leu  
 355 360 365  
 Ile Ala Thr Ile Ser Ser  
 370

## (2) INFORMATION FOR SEQ ID NO:151:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1777 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:151:

GGTCTTGACC ACCACCTGGG TGTCTAAGTC GGTGCCCCGA TTCAAGTCCA GGTACTCGTG 60  
 GGTGGGGTCG CGCAACCAAT ACCGACAAGC ATGCGAGCAG CCGCGGTAGC GTTGACCGCT 120  
 GTACCGAATC GGCACGCGG CCGGTTGGG CACCTTGTTC AGCGCTGATT TGCACAACAC 180  
 CTCGTGGGAG GTGATGCCGT CCAATTGTGG CCGCGCAAGC CTGCGGACCA GGCCTATCCG 240  
 CTGCAACCCG GCAGGCCCCG TGTGCAACCG GCATCCCTT CACCGGACG GCTTCCCCGG 300  
 CCGAACGCAAT ACCATTATTC GAAGAACCGT TCTATACCTT GTCAACGCTG GCGGCTACCG 360  
 AGCGGCGCAC AGGATGTGAT ATGCAATCTC TGCCGCGACA GACAGGAGCC AGGCTTATG 420  
 ACAGCAFTCG CGCTCGAGCC CTACGGGCG CCGAAGTACC TAGAATGCG CCGGAAGCGC 480  
 ATGGCGTATA TCGACGAGG CAGGGGTGAC GGCATCGTCT TTCAGCGCG CCAACCCACG 540

```

TCGTCTTACT TGTGGGSCAA CATCATGCCG CACTTGGAAg GGCTGGGCCG GCTGGTGGCC 600
TGGGATCTCA TCGGGATGGG GCGCTCGGAC AAGCTCAGCC CATCGGGACC CGACCGCTAT 660
AGCTATGGCG AGCAACGAGA CTTTTTGTTC GCGCTCTGGG ATCGGCTCGA CCTCGGCGAC 720
CACGTGGTAC TGGTCTGCA CCACTGGGGC TCGCGCTCG GCTTCGACTG GGCTAACCAG 780
CATCGCGACC GAGTGCAGGG GATCGCGTTC ATCGAAGCGA TCGTCACCCC GATGACGTGG 840
GCGGACTGGC CGCGGGGGGT GCGGGGTGTG TTCCAGGTTT TCGGATCGCC TCAAGGCGAG 900
CCAATGGCTT TGGAGCACA CACTTTTGTG GAACGGGTTC TGCGCGGGGC GATCCTGCGA 960
CAGCTCAGCG ACGAGGAAT GAACCACTAT CGCGGGCCAT TCCTGAACGG CGGCGAGGAC 1020
CGTCGCCCCA CGTTGTCTTG GCGACGAAAC CTTCCATCG ACGGTGAGCC CGCCGAGGTC 1080
GTGCGCTTGG TCAACGAGTA CCGGAGCTGG CTCGAGGAAA CCGACATGCC GAACTCTTC 1140
ATCAGCGCGG AGCGCGGCGC GATCATCAGC GCGCGCATCC GTCAGTATGT CAGGAGCTGG 1200
CCCAACCAGA CGAAATCAC AGTCCCGGTC GTGCATTTTC TTCAGGAGGA CAGCGATGGC 1260
GTGCTATCGT GGGCGGGCGC TCGGCAGCAT CGCGACCTG GGAGCGCTCT CATTTACGGA 1320
GACCAAGAAAT GTGATTTCCG GCGAAGGGGG CGCGTGTCTT CTCAGTCAAT AAGACTTCCT 1380
GCTCGGGCGA GAGATTCTCA GCGAAAAGGG CACCAATCGC AGCGGCTTCC TTCGCAACGA 1440
GGTCGACAAA TATAGCTGCC AGGACAAAGG TCTTCCTATT TCGCCAGCGA ATTAGTGGCT 1500
GCTTTTCTAT GGGCTCAGTT CGAGGAAGGC GAGCGGATCA CGCGTATCCG ATTGGACCTA 1560
TGGAACCGGT ATCATGAAAG CTTGGAATCA TTGGAACAGC GGGGCTCTCT GCGCGGTCCG 1620
ATCATCCAC AGGCTGCTC TCACAACGCG CACATGTAAT AGGTGTTACT AGCGCCGAGC 1680
GCGGATCGGG AGGAGGTGCT GCGCGGTCTG ACGAGCGAAG GTATAGGCGC GGTCTTTCAAT 1740
TACGTGCGGC TTCAGGATTC GCGCGCGGGG GGTGCT 1777

```

(2) INFORMATION FOR SEQ ID NO:152:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 324 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:152:

GAGATTGAAT CGTACCGGTG TCCTTAGCGG CTCGGTCCCG TGAATGCCCA TATCAGGCAC	60
GGCCATGTTT TGGCTGTGGA CCTTCGCCCC ATGCCCGGAC GTTGGTAAAC CCAGGGTTTG	120
ATCAATTAAT CCGGGGCAAG GTTCCGGGAA GCGGGCCAGG ATGTGCGTGA GCGGCCGCCG	180
CGCCGTGCGC CAGGGGACCG CTGGATGCTC AGCCCCGGTG CGGCGACGTA GCCAGCGTTT	240
GGCGCGTGTC GTCCACAGTG GTAATCCGGT GAGGACGGCG CCGGGTGCCT GGGTGAAGAC	300
CGTGACCGAC GCGGCCGATT CAGA	324

## (2) INFORMATION FOR SEQ ID NO:153:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1338 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:153:

GGGCTACCGC CGGTTCCCGC TGGCAGCGGA CCTGTACGAC CTGAACCACT TGGCTCCCG	60
AACGATTGAC GAACCGCTCG TCGGGGGGT GTGGGTGCTC AAGGTGTGGG GTGATGTCT	120
CGATGACCGG CGCGGCACCG GGGCACTAGC CCTCGAGAC GTCCTCGCCG CCGGCAGCGA	180
GCAGGACTTC CAGCCCGACT CGATCGGCGT GGTGACCCGT CCTGTGCTA TGGCTGCCG	240
GGAGGCTCGC GTTCGGAGGC GATTTGCGT CTCCTACTGAC CTCGACGCCG ACCAGCAGCG	300
GTGGGCGCGC TCGACCGAAC GGCACCGCG CGAGGTGAG AACCGGCTGG CGGTGCTCGC	360
GTCTGATCA ACCTGCCCGC GATCGTGCGG TTCCGCTGGC AGGTTGCGG CTGGACGCGG	420
CTGAATCGAC TAAATGAGAG CAGTTGGGCA CGAATCGGC TGTGCTGCTG AGCAAGACAC	480
GAGTACTGTC ATCACTATTG GATGCACTGG ATGACCGGCC TGATTCAGCA GGACCAATGG	540
AACTGCCCCG GGCAGAACGT CTCGGAGATG ATCGGGCTCC CCTCGGACC CTGCGGTGCT	600
GGGTCAATTC GGACATCGGT CCGGCTCGCG GCATCGTGGT GACGCCAGCG CTGAAGGAGT	660
CGAGGCGCGC GGTGCACCGC CTGCTGGAGC GCGGGCAGAC GGTGCTGCTG CTAAGGCGCG	720
GGATCGGCGA GAAGCGCTTC GAGGTGCGG CCGACGATT CTTGTTGTTT CCGACGGTCC	780
CGCACAGCCA CGCGGAGCGG GTTCGCCCCG AGCACCGCGA CCTGCTGCGC CCGGCGGCGG	840

```

CCGACAGCAC CGACGAGTGT GTGCTACTGC GGGCCGACGC GAAAGTTGTT GCGGCACATGC      900
CGGTTAACCg GCCAGAGGGT CTGGACGCCA TCGAGGATCT CCACATCTGG ACCGCCGAGT      960
CGGTGCGGCG CGACCGGCTC GACTTTCGGC CCAAGCACAa ACTGGCCGTC TTGGTGGTCT      1020
CGGCGATCCC GCTGCCCCGAG CCGGTCCGCG TGGCGGTAG GCGCGATAC GCGGCTTCCA      1080
CCAGCTGGGT GCACTCGCG GTGACGCCGA CGTTGGCGGC GCGGCTGCAC GACGAGGCCG      1140
CGCTGGCGGA GGTGCGCGCC CGGCTCCGCG AGGCGGTGGG TTGACTGGGC GGCATCGCTT      1200
GGGTCTGAGC TGTAGCCCA GTGCGGCTC CGAGTGAATC GCTGTCGTT CGGTCCCTGC      1260
TGGCGTCAT TGAAGGCGCG GGCACAGCA GCATTGGCG GCGCATCTC CCGCGCGCGG      1320
GGGCGCACCG CTACAACC                                     1338

```

## (2) INFORMATION FOR SEQ ID NO:154:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 321 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:154:

```

CGGGCGGGCAC CGGCGGCACC GCGGTACCG CGGCGACCG CGCTGACGCC GCTGCTGTGG      60
TGGGCTTCCG CGCGAACGGC GACCTTGGCT TGGTGGCG CAAGGGGCTT AACGCGGGA      120
TAGCTGGGCG CGCGCTGACA GCGGGGCTCG CGGCGACCG CGGCACCGCG GCGAAGCTC      180
GCACCGGCGG TCGCGCGGCG GCGGCAACG AGCGCGGCG CACCGGCAAT CCGCGCGGTA      240
AGGGCGGCGA CGGCGGATC GCGGTGCGG GCGGGGCGCG CGGCGCGGCG GCGACCGGCA      300
ACGCGGCGCA TCGCGGCAAC C                                     321

```

## (2) INFORMATION FOR SEQ ID NO:155:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 492 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:155:

```

GAAGAGCCCGG CCCC GCCATA TCGATCGGCT CGCGGACTAC TTTCGGCEAA CGTGCACGGG      60
GGGGCGTCCG GCTGATCATC ACCGGTGGCT ACGCGCCCA CCGCACCGGA TGAGTGTCTGC      120
CGTTGGGCTC CGAACTCGTC ACTTCGGCGC AAGCCCGAGC GCACCGCCGA ATCACCAGGG      180
CGGTCCACGA TTCGGGTGCA AAGATCGTGC TGCAAATCCT GCACGCGGA CGCTACGGCT      240
ACGACCCACT TCGGTTCAGC GCTCGCCGA TCAGGCGGC GATCAGCCCG TTTCGTCCGC      300
GAGCACTATC GCTTCGCGGG GTCGAAGCGA CCATGCGCA TTTCGCGCG TCGCCCGAGT      360
TGGCCCGCGA TCGGGCTAC GACGGCGTCG AATCATGGG CAGCGAAGGG TATCTGCTCA      420
ATCAGTTCTT GCGCGCGCG ACCAACAAGC GCACCGACTC GTGGGCGCGC ACACCGGCGA      480
ACCTCGCGCG GT                                         492

```

## (2) INFORMATION FOR SEQ ID NO:156:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 536 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:156:

```

Phe Ala Gln His Leu Val Glu Gly Asp Ala Val Glu Leu Trp Arg Ala
1           5           10           15
Asn Ala Ala Asp Gln Ala Asp Pro Leu Gln Pro Gly Ser Ala Arg Arg
20          25          30
Gln Arg Ala Ser Arg Ser Pro Arg Arg Leu Ala Gly Pro Asn Ala Tyr
35          40          45
His Tyr Ser Asn Asn Arg Ser Ile Leu Cys Gln Arg Trp Pro Leu Pro
50          55          60
Ser Ala Ala Gln Asp Val Ile Cys His Leu Cys Pro His Arg Gln Glu
65          70          75          80
Pro Gly Leu Met Thr Ala Phe Gly Val Glu Pro Tyr Gly Gln Pro Lys
85          90          95
Tyr Leu Glu Ile Ala Gly Lys Arg Met Ala Tyr Ile Asp Glu Gly Lys

```

174

	100		105		110														
Gly	Asp	Ala	Ile	Val	Phe	Gln	His	Gly	Asn	Pro	Thr	Ser	Ser	Tyr	Leu				
	115						120					125							
Trp	Arg	Asn	Ile	Met	Pro	His	Leu	Gln	Gly	Leu	Gly	Arg	Leu	Val	Ala				
	130					135					140								
Cys	Asp	Leu	Ile	Gly	Met	Gly	Ala	Ser	Asp	Lys	Leu	Ser	Pro	Ser	Gly				
145					150					155					160				
Pro	Asp	Arg	Tyr	Ser	Tyr	Gly	Gln	Gln	Arg	Asp	Phe	Leu	Phe	Ala	Leu				
				165					170					175					
Trp	Asp	Ala	Leu	Asp	Leu	Gly	Asp	His	Val	Val	Leu	Val	Leu	His	Asp				
			180					185					190						
Trp	Gly	Ser	Ala	Leu	Gly	Phe	Asp	Trp	Ala	Asn	Gln	His	Arg	Asp	Arg				
	195						200				205								
Val	Gln	Gly	Ile	Ala	Phe	Met	Gln	Ala	Ile	Val	Thr	Pro	Met	Thr	Trp				
	210					215					220								
Ala	Asp	Trp	Pro	Pro	Ala	Val	Arg	Gly	Val	Phe	Gln	Gly	Phe	Arg	Ser				
225					230					235					240				
Pro	Gln	Gly	Gln	Pro	Met	Ala	Leu	Gln	His	Asn	Ile	Phe	Val	Gln	Arg				
				245					250					255					
Val	Leu	Pro	Gly	Ala	Ile	Leu	Arg	Gln	Leu	Ser	Asp	Gln	Gln	Met	Asn				
			260					265					270						
His	Tyr	Arg	Arg	Pro	Phe	Val	Asn	Gly	Gly	Gln	Asp	Arg	Arg	Pro	Thr				
	275						280					285							
Leu	Ser	Trp	Pro	Arg	Asn	Leu	Pro	Ile	Asp	Gly	Gln	Pro	Ala	Gln	Val				
	290					295				300									
Val	Ala	Leu	Val	Asn	Gln	Tyr	Arg	Ser	Trp	Leu	Gln	Gln	Thr	Asp	Met				
305					310					315					320				
Pro	Lys	Leu	Phe	Ile	Asn	Ala	Gln	Pro	Gly	Ala	Ile	Ile	Thr	Gly	Arg				
			325						330					335					
Ile	Arg	Asp	Tyr	Val	Arg	Ser	Trp	Pro	Asn	Gln	Thr	Gln	Ile	Thr	Val				
			340					345					350						
Pro	Gly	Val	His	Phe	Val	Gln	Gln	Asp	Ser	Asp	Gly	Val	Val	Ser	Trp				
	355						360					365							
Ala	Gly	Ala	Arg	Gln	His	Arg	Arg	Pro	Gly	Ser	Ala	Leu	Ile	Ser	Arg				
	370					375					380								
Asp	Gln	Gln	Cys	Asp	Phe	Arg	Arg	Arg	Arg	Arg	Pro	Ala	Cys	Gln	Leu				
385					390					395					400				

```

Ile Arg Leu Pro Ala Pro Gly Arg Asp Ser Gln Gly Lys Gly His Gln
      405                      410                      415

Ser Gln Pro Leu Pro Ser Gln Arg Gly Arg Gln Ile Tyr Val Ala Gly
      420                      425                      430

Gln Arg Ser Ser Tyr Leu Pro Ser Gln Leu Val Ala Ala Phe Leu Trp
      435                      440                      445

Ala Gln Phe Glu Glu Ala Glu Arg Ile Thr Arg Ile Arg Leu Asp Leu
      450                      455                      460

Trp Asn Arg Tyr His Gln Ser Phe Glu Ser Leu Glu Gln Arg Gly Leu
      465                      470                      475                      480

Leu Arg Arg Pro Ile Ile Pro Gln Gly Cys Ser His Asn Ala His Met
      485                      490                      495

Tyr Tyr Val Leu Leu Ala Pro Ser Ala Asp Arg Glu Glu Val Leu Ala
      500                      505                      510

Arg Leu Thr Ser Glu Gly Ile Gly Ala Val Phe His Tyr Val Pro Leu
      515                      520                      525

His Asp Ser Pro Ala Gly Arg Arg
      530                      535

```

## (2) INFORMATION FOR SEQ ID NO:157:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 264 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:157:

```

Asn Glu Ser Ala Pro Arg Ser Pro Met Leu Pro Ser Ala Arg Pro Arg
1           5           10           15

Tyr Asp Ala Ile Ala Val Leu Leu Asn Glu Met His Ala Gly His Cys
      20           25           30

Asp Phe Gly Leu Val Gly Pro Ala Pro Asp Ile Val Thr Asp Ala Ala
      35           40           45

Gly Asp Asp Arg Ala Gly Leu Gly Val Asp Glu Gln Phe Arg His Val
      50           55           60

Gly Phe Leu Gln Pro Ala Pro Val Leu Val Asp Gln Arg Asp Asp Leu

```



176

65		70		75		80
Gly Gly Leu Thr Val Asp Trp Lys Val Ser Trp Pro Arg Gln Arg Gly						
	85			90		95
Ala Thr Val Leu Ala Ala Val His Glu Trp Pro Pro Ile Val Val His						
	100			105		110
Phe Leu Val Ala Glu Leu Ser Glu Asp Arg Pro Gly Gln His Pro Phe						
	115			120		125
Asp Lys Asp Val Val Leu Gln Arg His Trp Leu Ala Leu Arg Arg Ser						
	130			135		140
Glu Thr Leu Glu His Thr Pro His Gly Arg Arg Pro Val Arg Pro Arg						
	145			150		155
His Arg Gly Asp Asp Arg Phe His Glu Arg Asp Pro Leu His Ser Val						
	165			170		175
Ala Met Leu Val Ser Pro Val Glu Ala Glu Arg Arg Ala Pro Val Val						
	180			185		190
Gln His Glu Tyr His Val Val Ala Glu Val Glu Arg Ile Pro Glu Arg						
	195			200		205
Glu Gln Lys Val Ser Leu Leu Ala Ile Ala Ile Ala Val Gly Ser Arg						
	210			215		220
Trp Ala Glu Leu Val Arg Arg Ala His Pro Asp Gln Ile Ala Gly His						
	225			230		235
Gln Pro Ala Gln Pro Phe Gln Val Arg His Asp Val Ala Pro Gln Val						
	245			250		255
Arg Arg Arg Gly Val Ala Val Leu Lys Asp Asp Gly Val Thr Leu Ala						
	260			265		270
Phe Val Asp Ile Arg His Ala Leu Pro Gly Asp Phe						
	275			280		

(2) INFORMATION FOR SEQ ID NO:158:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 264 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:158:

ATGACATCT	CTCCGGTGGT	GGTCCCAAG	GCCTTTGCG	GATTCGCCGG	CTACTCCTCC	60
GCCATGCACG	CAATCGCCCG	TTTCTCCGAT	GGTTGCGCC	AAGAGCTGCG	GGGTAGCGGA	120
ATCCCGCTCT	CGGTGATCCA	CCCGGCGCTG	ACCCAGACAC	CGCTGTTGGC	CAACCTCGAC	180
CCCGCCGACA	TGCCGCGCCG	GTTCGCGACG	CTCAGGCCCA	TTCCCGTTCA	CTGGGTCCGG	240
GCACCGGTGC	TTGACGCTCT	GGCG				264

(2) INFORMATION FOR SEQ ID NO:159:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1171 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:159:

TAGTCGGCGA	CGATGACGTC	GGGTCCAGG	CCGACCGCTT	CAAGCAACCAG	CGCGAACCAG	60
AAGCCGGTGC	GATCCTTACC	CGCGAAGCAG	TGGGTGAGCA	CCGGGCGTCC	GGCGGCAAGC	120
AGTGTGACGA	CACGATGTAG	CGCGGCTCT	GCTCCATTCG	GCCTTGGGAA	TTGGCGATAC	180
TCGTGGGTCA	TGTAGCGGGT	GGCGGCTCA	TTTATCGACT	GGCTGGATTC	GCCGACTCCG	240
CGTTTGACCC	CTTCATTGGT	TAGCAGGCTC	TTGAATGCGG	TTTGGTTCGG	CGCTGAGTCC	300
TGGCGCTCAT	GATCGGCGAG	GTGGGGGAG	GGCAGCGAGT	GGACGTCCAT	GGCGTCCGGA	360
ACCCGCTCTG	GACCGCGCGG	GGCAAGCTTC	CGGAGCGACC	GCAGGTCCGG	AACGTCCGTC	420
ATCCCCAGCC	GGCGCAGCGT	TGCCCTCTCT	GCCGATTTCC	GCACGAGGCT	GGCGAGCCAC	480
CGGGCATCAC	CAAGCAAGCC	TTCCCCAGTA	CGGATCTTCA	CTTCGCGATC	CGGCAGACCA	540
ATCTCCTCGC	CGCCCATCGT	CAGATCCGCG	TCGTGCGTTG	ACAAGAACGG	CGGCAGATCT	600
GCCAGCGGGT	ATCGGAGATT	GAACCGCGCA	CGCAGTTCTT	CAATCGCTTC	GCGCTGCCGC	660
ACTATTGSCA	CTTCCGGCG	GTGCGGTAT	TCAGCAAGCA	TGCGAGTCTC	GACGAACTCC	720
CUXACGTAA	CCACCGCGT	AGCTCCCGGC	GTGACGCGGA	CGATCGGCGG	GTGATCTTTC	780
CGGCCAGCGT	CTAGCCGCTT	GATCCAGCGC	TTCCGCGTTC	CGGCGGGGAG	GCCGATCAGC	840
TTATCGACCT	CGGCGTATCC	CGACGCGAAG	CTGGGCGGCT	TGCTCGAGGT	CAAGAACTCC	900
ACCATCGGCA	CGGCACCAAA	GGTCCCGCAC	CTGACTTACG	TGCGCGAGCG	CGACATCGGC	960

```

GAGTACAGCA ACATCGGGGC CTCAGCGTG TCGTCAACT ACCAGGTAC GTCCAAACGG      1020
GGCACCACCG TCGTTCCGA CTACGGACC GGGTCGGACA CCATGTTTGT GGGCCCASTA      1080
ACCATCGGCG ACCGCGGTA TACGGGGCC GGCACATGG TCGGGAGGA TGTCCCGCC      1140
GGGGCGCTGG CAGTGTGGC GGSTCCCAA C                               1171

```

## (2) INFORMATION FOR SEQ ID NO:160:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 227 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:160:

```

GCAAAGGCGG CACCGGCGGG GCGGCGATGA ACAGCCTCGA CCGGCTGCTA GCGGCCCAAG      60
ACGGCGGCCA AGGGGACACC GCGGCGACCG GCGGCGACCG CCGGCGCGGC GGCACCGTCT      120
TCACCCAAGG CCGCGACCGC AACCGCGCCA ACCGCGGTGA CCGCGGGGTC GCGGCGCAAG      180
GCGGAACCGG CGGAAACGGC GCAGACAACA CTACCGCGCG CCGCGCC                               227

```

## (2) INFORMATION FOR SEQ ID NO:161:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 304 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:161:

```

CCTCGCCACC ATGGGCGGGC AGGGCGGTAG CGGTGCGGCC GGCTCTACCC CAGGCGCCAA      60
GGGCGGCCAC GCGTTCACTC CAGCCAGCGG CCGCGACCGC GCGGACGCGG GCAACGGCGG      120
CAACTCCCAA GTGGTCGGCG GCAACGGCGG CAGCGCGGCG AATGGCGGCA ACGGCGGCGG      180
CGCGGGCAGG GCGGGCAACG GCGGCGCGCG CCGCGACCGC GCGTTGCTG GCATGAGTGC      240
CAACGCCACC AACCTGGTG AAGACGCGCC AAGCGGTAAC CCGGCGGCGA ACGGTGGCGC      300

```

CGGC

354

(2) INFORMATION FOR SEQ ID NO:162:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1439 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:162:

```

GTGGGAGGCT GCGGAGGCTG TATAACAAGG ACRACATCGA CCAGCGCCCG CTGGGTGAGC      60
TGATCGACCT ATTTAACAGT GCGGCGTTCA GCGGGCAGGG CGAGCAGCGC GCGCGGGATC      120
TGATGGCTCA GGTCTACGAA TACTTCTCG GCAATTTGGC TCSCGCGGAA GGGAGCGGG      180
GTGCGGAGTT CTTTACCCCG CCCAGCGTGG TCAAGGTGAT CGTGGAGGTG CTGGAGCCCT      240
CGAGTGGGCG GGTGTATGAC CCGTCTCGG GTTCCCGAGG CATGTTTGTG CAGACCGACA      300
AGTTCATCTA CGAACAGGAC GCGGATCGGA AGGATGTCTC GATCTATGGC CAGGAAAGCA      360
TTGAGGAGAC CTGGCGGATG GCGAGATGA ACCTCGCCAT CCACGCGATC GACAACAAGG      420
GGTTCGGGCG CCGATGGAGT GATACCTTCG CCGCGGACCA GCACCCGAC GTGCAGATGG      480
ACTACGTGAT GCGCAATCG CCGTTCAACA TCAAAGACTG GCGCCGCAAC GAGGAAGACC      540
CAGCGTGGCG CTTCGGTGTT CCGCCCGCCA ATACCGCCAA CTACCGATCG ATTACGACA      600
TCCTGTACAA CTTCGCGCGG GAGGTCGGG CCGGCTGGT GATGCGCAAC GGTCGATGT      660
CGTUGAACTC CAACGGCAG GGGGATATC GCGCGCAAT CGTGGAGGCG GATTIGSTTT      720
CGTGCATGGT CGCTTACCC ACCAGCTGT TCGCGAGCAC CGGAATCCCG GTGTGCTGT      780
CGTTTTTCCG CAAAACAGG GCGGCAGGTA AGCAAGGCTC TATCAACCG TCGGGGCGAG      840
TGCTGTTCAT CGACGCTCGT GAACTGGGCG ACCTAGTGA CCGGGCCGAG CCGGCGCTGA      900
CGACAGAGGA GATCGTCCG ATCGGGGATA CCTTCCACCG GAGCAGGACT ACGGCGACG      960
CGGCTCGCG TGCTCCCGG GGTAAATGGG GCACTGGCT CAACCGCGCG GCGGTGCTG      1020
GCGGGGCGCG CCGCAACCG GGTGTGCGCG GGTGTGCTT CGGCAACCT GTGGGCGCG      1080
ACGGCGGCGA CGGCGGCAAC GCGGGCGAG GCGGCGACG CAGGACGGG GCGGCGGCG      1140
GCAAGGGCGG CAACGGCAGC AGCGGTGCG CCGCGGCTC AGCGGTGCTC AACGTGCTG      1200

```

CGGGCCACGG	CGGCNAAGGC	GGCAATGGCG	GCAACGGCGG	CAACGGCTCC	GGGGGTGCTG	1260
GGGGCCAGGG	CGGTGCTGGC	GGCAGCGCGG	GCAACGGCGG	CCACGGCGGC	GGTGCCACCG	1320
GGGGCGCCAG	CGGCNAAGGC	GGCAACGGCA	CCAGCGGTGC	CGCCAGCGGC	TCAGGGGTCA	1380
TCACGTTCAC	CGCGGGCCAC	GGCGGCAACG	GGGGCAATGG	CGGCAACGGC	GGCAACGGC	1439

## (2) INFORMATION FOR SEQ ID NO:163:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 329 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:163:

GGGCGGGCGG	GGCGGATTT	TCTGCTGCT	TCATTGTGCG	TGGGATAAC	GGCGTGATG	60
GTGCTAACGG	CGGATGGGC	GGGCTGGCG	GGGCTGGCGG	CCCCGGGGG	GGGCGGGCC	120
TCATCAGCT	GCTGGCGGG	CAAGCGCGCG	GGGGGGCGCG	CGGACCGGC	GGGCGGGCG	180
GTGTGGCGG	TCAGGGGGG	GGGGGGGGC	GGGCAACCA	GGCTTCAAC	GCAGGTGCGG	240
GGGGGGGGCG	CGGCTGATC	AGCTGCTGG	GGGGCGAGG	CGCGGGCGGG	GGGGGGGGGA	300
CGGCGGGGGC	CGGCGGTCTT	GGCGGTGAC				329

## (2) INFORMATION FOR SEQ ID NO:164:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 80 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:164:

GCAACGGTGG	CACGGCGGC	ACGAGCAGCA	CCGTGGGGAT	GGCGGGAGGT	AAGTGTGCTG	60
CGCGCGGGCT	GATCGGCAAC					80

## (2) INFORMATION FOR SEQ ID NO:165:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 392 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:165:

```

GGGCTGTGTG GCATTCACAC CGCCGCAATC GCGGACGTTG GCGGCCAAT ATCCAGCTCA      60
AGGCCTACTA CTACCGCTCG GAGGACCGCC GCATCAAGGT GCGGCTCAGC GCGCAAGGAA      120
TCAAGGTCAAT CGACCCCGAC GGGCATCGAG GCGCTCCTCG CCGGCTTCGG GCAGGATCCG      180
CCCCGGCGCA CTTCGGCGGC CAAGCGGGCT CATCGCTCGG AACGGCGGGG ATCCTGTGAG      240
CACAACTGAT GCGGCCCAAC GAGATTCGTC CAATTGTCAA GCGGTGTTCG ACCGCAGGGA      300
CGGGTTATAC GTATGTCAAC CTATGTCACT CGCAGCAACC GGCATAACGA TCCCGTGATC      360
CGCGGACAGC CCACCACTCC AAGACCGTTA CA                               392

```

## (2) INFORMATION FOR SEQ ID NO:166:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 535 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:166:

```

ACCGGCGCCA CCGGCGGCAC CGGCTTCGCT GGTGGCGCGG GCGGGGCGGG CCGGCAGGGC      60
GGTATCAGCG GTGCGGGCGG CACCAACGCG TGTGGTGGCG CTGGCGGCAC CCGCGGACAA      120
CGCGGCGCGG GGGGCTCTGG CGGCGCGGCG GCGGATAACC CCACCGGCAAT CCGGCGGCGG      180
GCGGGGACCG GCGGCACCGG CGGAGCGGCG GAGGCGGCGG GGGGCGGTGG CCGCATCGGT      240
ACCGGCGGCA CCGGCGGCGG GGTGGGCGCG GTGCGTAAGG CCGGGATCGG CGGTACCGGC      300
GGTACGGGTG GTGTGGGTGG TGCTGGTGGT GCAGGTGCGG CTGCGGCGCG TGCGACGAGC      360
GCTACCGGTG GCGCGGGGTT CGCGGCGGCG GCGGCGGCGG AAGGCGGAGC GCGCGGCAAC      420
AGCGGTGTGG GCGGCACCAA CGGCTCGGCT GCGCGCGGCG GTGCAAGCGG CAAGCGGCGG      480

```

ACCGGAGGTT CUGGGGGGTC CGGCGCGGAC AACCCACCGG GTGCTGGTTT CCGCG 535

(2) INFORMATION FOR SEQ ID NO:167:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 690 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:167:

CCGACGTCC CCGGGCGATA CCGGGGTAC CGACTACTAC ATCATCGCA CCGAGAATCG	60
GCGGTGCTC CACCGCTCC GGGCGGTCC GGTATCGCA CATCGGCTCG CCGACCTGAT	120
CGAGCGAAC CTGAAGGTCA TGTCAACT GGGCTACGC GACCCGACT ACGGCTACTC	180
GACGAGCTAC GCGATGTGC GAGCGCGTT GGGCTGTGG CCGACGTGC CGCCTCAGGT	240
CATCGCGAT GCGCTGCCC CCGGAACCA AGAAGGATC CTGACTTCA CCGCCGACCT	300
GGGCGCTC TCGCGCAAC CGTCAAGCT CCGCGATC CAGCTGCCG AACCGCGCA	360
TCTGCTGCC GCGGTGCGG CCGCACCGAC GCGGCGGAG GTGTGAACA CGCTCGCCAG	420
GATCATCTC ACCAACTAC CGGTCTGCT GCGACCGTG GACATCGCC TCGCTGCTC	480
ACCAAGCTC CGCTGTAC CCGCAACTG TTGCTCAGC AACTCGCTC GGGCAATCTC	540
ATCAAGCGA TCGCTATCC CTGCGCGCC ACCTAGGTT TAGGCAGAT CGATAGCGG	600
CGGCTGGA TGTCAACC TCTCGCGG GCGCTCGAC ACGTTGAA ACATCGAGG	660
CGTCTCAC TAACGATC CCGACGGAT	690

(2) INFORMATION FOR SEQ ID NO:168:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 407 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:168:

```

ACGCTACCG CGGTACTGGC GCGCGCCACG GCGGCAACCG CCGGAATCCC GGGTGGCTCT      60
TGGGCACAGC CCGGCGTGGC GGCACCGGTG GCGCCCGCAG CACCGGTACT GCAGGTGGCG      120
GCTCTGGGGG CACCGGGGGC GACGGGGGGA CCGCGGGGGG TGGCGGCTTG TTAATGGGGG      180
CGGGCGCCCG CCGGCAAGGT GCACTGGCG GCGCGGGCGG TCGCGGTGTC GACGCTGGCG      240
GCGCCGGCGG GCGCGGGGGG GCGCGCGGCA ACCGCGGGCG CCGCGGTCAA GCGCGCTGTC      300
TCTTCGGGGG CCGCGGCACG GCGCGAGCGG GCGGCTACGG CCGCGATGGC GGTGGGGGGG      360
GTGACGGCTT CGACGGCAGC ATGCGCGGCG TGGGTGGTAC CGGTGGC      407

```

{2} INFORMATION FOR SEQ ID NO:168:

- {1} SEQUENCE CHARACTERISTICS:
  - {A} LENGTH: 468 base pairs
  - {B} TYPE: nucleic acid
  - {C} STRANDEDNESS: single
  - {D} TOPOLOGY: linear

{x1} SEQUENCE DESCRIPTION: SEQ ID NO:169:

```

GATCGGTGAG CAGATCGGCC TCGCGGGCAA CGGATTGCGC GGTCTCACC GAGAACATCG      60
TGCAGCGGGC GCGCGGAGC AGCCGCTGC GCTCGGGGCG GTGGAACGCC TCGAGCAGGC      120
ACAGCCAGTC CTTGGCGGGC TCGAGGGCGA ACAGCTGGGT CTCACCGGTG TAGATCGGCG      180
GGATGCCCCG CTCGCCAAC GCATTCGGGC ACGCCCGGCG GTCTTTGTGA TGCCTCAGCA      240
TCACCGCGAT GTCTCGGCC ACCAGGGGCC GCGCGGGGAA GGTGGCCCCG CTGGCCAGTA      300
GCGCGCGGAC GTCGCGGCC AGGTGCTCGG GATGTGCGG GCGCAGCGCT CCGCGCGGAC      360
GCGCGAAAAA CAGCCCTCA CCGAGCTGGG TCGCGCTGGC ATATCCCTTG CGGTCTGGG      420
CGATATTGGA CGCCCATGCC CGACCGCGT ACAGGCGGGC CAGCACG      469

```

{2} INFORMATION FOR SEQ ID NO:170:

- {1} SEQUENCE CHARACTERISTICS:
  - {A} LENGTH: 219 base pairs
  - {B} TYPE: nucleic acid
  - {C} STRANDEDNESS: single
  - {D} TOPOLOGY: linear



(x1) SEQUENCE DESCRIPTION: SEQ ID NO:170:

```

GCTGGTAACG GCGGCCACGG TGGCATCGGC GCGCGCGCGG ACAGAGGCGC CGACGGCGGC      60
GGCCCCAATG CTAAACGGCC AAACGGCGAG AACGGCGGTA GCGGTGGTAA CGGTGGCGAC      120
GGCGGCGCGG GCGGCAATGG GGGGCGGGGC GCGACGGCGC AGGCGGCGCG GTACACCGAC      180
GGCGGCGCGG GCACCGGCGG CGACGGCGGC AACGGCGGC      219

```

(2) INFORMATION FOR SEQ ID NO:171:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 494 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:171:

```

TAGCTCGGCG GAGGGGCGCA AGGGCGGCGA CGGTGGCCAC GCGGCTGACG GGTGCGCGCG      60
CAACAGTTTC GTACCCCAAG GCGGCAGCGG CGGTGGCGGC GCGGCGCGCG GCGGCGCGCG      120
CAGCGGCTTT TTGCGCGCGA AGGGCGGCTT GCGGCGCGAC GCGGCTCAGG GCGGCGCGCA      180
CGCGGCGCGT ACCGTCGCGA CGGTGGCGCG TGGCGGCGGC AACGCGGCTG TCGGCGCGCG      240
GGGCGGCGAC GCGGTCTTTG CGGTGGCGCG GCGGCAGGCG GCGGCTCGGTG GCGGCGCGCG      300
CAATGGCGGC GGTTCACCG GCGGCACCGG CGGCTTGGCG GCGGCGCGCG GTGGCGCGAG      360
CAACGCGCGG GTGCTGCGG AATCGGCGCT GACCATGGAC AGCGCGCGCA AGTTGCGTGC      420
CATCGCATCA GCGCGTACT GCGCGAACA CCGGAACAT CACCGAGTT AGCGGGCGCG      480
ATTCTCTGAT CAGC      494

```

(2) INFORMATION FOR SEQ ID NO:172:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 220 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:172:

GGGCGGGTGG TGGCGCGGGC CAGCTCTTCA GCGCGGGAGG GCGCGCGGGT GCGGTTGGGG	60
TTGGCEGCAC CGGCGGCCAG GGTGGGGCTG GCGGTGCGGG AGCGGCGGGC GCGGACGCCC	120
CCCCCAGCAC AGGTCTAACC GGTGCTACCG GGTTCGCTGG CCGGGCGGGC GCGGTCGGCG	180
GCGAGAGCGG CAACGCCATT GCGGCGGGCA TCAAGGGCTC	220

## (2) INFORMATION FOR SEQ ID NO:173:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 388 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:173:

ATGGCGGCAG CCGGGGCCCC GCGGCTGCTG GCGGGGCCCC CGACTACAT TTCCACGGC	60
GGGCAGGGTG GTGCGGGGGG CCAAGCGGGC CAGGCGGGC TGGGCGGGC AAGCACCAC	120
TGATCGGCCT AGCGGCACCC GGGAAAGCCG ATCCAACAGG CGACGATGCC GCCTTCCTTG	180
CCGCGTTGGA CCAGGCGGGC ATCACCCTAG CTGACCCAGG CCACGCCATA ACGGCGGCCA	240
AGCGGATGTC TGGGCTGTGT GCTAACGGCG TAACAGGTCT ACAGCTGGTC GCGGACCTGC	300
GGGACTACAA TCCGGGCTG ACCATGGACA GCGCGGCCAA GTTCGCTGCC ATCGCATCAS	360
GCGGCTACTG CCGGAAACAC CTGGAACA	388

## (2) INFORMATION FOR SEQ ID NO:174:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:174:

GCAAGCGCGG CACCGCGGGG GCGGGCATGA ACAGCTTCGA CCGGCTGCTA GCGGCCCCAG	60
ACGGCGGCCA AGGCGGCACC GCGGCGACCG GCGGACAGCG CGGCGCGGGC GGCACCAAGT	120

186

```

TCACCCAAAGG CCCCCACGGC AACGCTCGCA ACGGCGGTGA CCGCGGGGTC GGCAGCAACG      180
GCGGAAACGG CCGAAACGGC GCAGACAACA CCACCACGGC CCGCGCCGGC ACCACAGGCG      240
GCGACGGCGG GCGCGCGGGG GCGGCGGAA CCGGCGGAAC CCGCGGAGCC GCGGCGACCG      300
GCACCGGCGG CCAACAGGC AACGCGCGCA ACGGCGGCAC CCGCGGCAAA GCGGCGACCG      360
GCGGCGACCG TCGACTCTCA GCGAGCACCG CTGCTGCGCG      400

```

(2) INFORMATION FOR SEQ ID NO:175:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 538 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:175:

```

GGCAACGGCG GCAACGGCGG CATCGCCGGC ATTGGCGGGC AACGGCGTTC CCGGACGGGG      60
AGCGGCAACG GCGGCCAAGC GCGGCAGCGG CCGCAACGGC GCGAAGCGCG GCATGGGCGG      120
CAACAGCGGC ACCGCGACCG GCGACGGCGG TCGCGCGGG AACGCGCGCG CCGCGGGCAC      180
GGGCGGCACC GCGGCGGACG GCGGCTCAC CGGTACTGGC GGCACCGGGC GCAAGCGTGG      240
CAGCGCGCGT GAGCGCGGTA ACGCGCGCAA CCGAGCAGAT AACACGGCAA ACATGACTGC      300
GCAGCGGCGC GGTACGGTG GCAACGGCGG CAGCGTGGC TTCGCGCGCG GCGCGGGGGC      360
CGGCGCGCGT CGCTTGACCG CTGCGCGCAA CCGCACCGGC GCGCAGGCG GCGCGGCGCG      420
CGATGCGCGC AACCGGGCCA TCGCGGGCCA CCGCCCTCTC ACTGACGACC CCGCGGCGAA      480
CGGGGCGACC GCGGCCAAGC GCGGCACCGG CCGCACCGGC GCGCGGGCA TCGGCAGC      538

```

(2) INFORMATION FOR SEQ ID NO:176:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 239 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:176:

GGGCGCGTGG	TCCCCGCGGC	CAGCTCTTCA	GCGCCGGAGG	CGCGGCGGGT	GCCTTGGGG	60
TTGGGGGCAC	CGCGGCCAG	GGTGGGCTG	GGGTGCGCG	AGCGGCGGG	GCCGACGCCC	120
CGCCAGCAC	AGTCTAAC	GGTGTACCG	GGTTCCTGG	CGGGGCGGG	GGCTTCGGG	180
GCCACGGGG	CAACGCCATT	GGGGCGGCA	TCAACGGCT	CGGTGGTGG	GGCGGCACC	239

## (2) INFORMATION FOR SEQ ID NO:177:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 985 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:177:

AGCAGCGCTA	CGGTGGGCG	CGGCTTCGCG	GGCGGCGCG	CGCGAGAGG	CGGAGCGGG	60
GGCAACAGCG	GTGTGGGCG	CACCAACGCG	TCCGCGGCG	CGCGCGGTG	AGCGGCAAG	120
GGCGGCGCG	GAGGTGCGG	CGGTTCGCG	CGGACAGG	CGACCGGTG	TGGTTTCGG	180
GGTGGCGCG	CGCGCACAG	TGGCGGCGG	GGCGGCGCG	GGCGGCGCG	GGCGACCGG	240
ACCGGCGCG	CGCGCGGCG	TGTGGCGCG	ACCGGTAGT	CAGGCGTGG	CGCGGCGCG	300
GGCGGCGCG	GTGACGGCG	CGATGGGCG	AGCGGTCTG	GGCTGGGCG	CTCGGCGTT	360
GACGGCGCG	AAGCGGCG	AGCGGCGCG	GGCGGCGCG	CGCGGCGCG	CGCGATCAAC	420
GGCGGCGCG	GGCGGCGCG	CAACGGCGG	GACGCGGCG	ACGGGCGAAC	CGGTGCGCG	480
GGTCTCGCG	ACAACGGCG	GGTGGCGGT	GACGTTGGG	CGGTGCGCG	CGCGGCGAAC	540
GGCGGCGAG	CGGGCGTGG	CCTGACAGG	AAGGCGGCG	ACGGGCGGCG	CGCGGCGAAT	600
GGCGGCGAAC	CGGGCGGCG	CGGTGCTGG	GGGCGGCGG	ACAACAATTT	CAACGGCGG	660
CAGGTTGGTG	CGGGCGGCG	AGCGGCGCAA	GGCGGCTTG	GGGGGCGAAG	CGCCACCTGA	720
TGGGCTTAGC	CGCACCGGG	AAGCGCGATC	CAACAGGCG	CGATGCGGCG	TTCCTTGGCG	780
CGTTGGACCA	GGCGGCGATC	ACCTAGGCTG	ACCGGCGCG	CGGTAAGCG	GGCGGCGAAG	840
CGATGTGTGG	GCTGTGTGCT	AACGCGGTA	CAGGCTTACA	GCTGCTGCG	GACCTGCGGG	900
AATACATTC	CGGGCTGACC	ATGACAGCG	CGCGCAACTT	CGCTGCGATC	GCATCAGGCG	960

CCTACTGCCC CGAACACCTG GAACA

985

(2) INFORMATION FOR SEQ ID NO:178:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2136 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:178:

CGGCACGAGG ATCGGTACUU CGCGGCATCG GCAGCTGCCG ATTCCCGGGG TTTCGCCACC	60
CGAGGAAGG CGCTACCAA TGGGCTGCGC GAAGTAGGGC GATCGCTTCG CGATGCCGGC	120
ATGAACGGGC GGCAICAAAT TASTCAGGA ACCTTTCACT TTAGCGACGA TAATGGCTAT	180
AGCCTAAGG AGGATGATCC GATATGACGC AGTCGCAGAC CCGACCGGTG GATCAGCAAG	240
AGATTTTGAA CAGGCCCCAAC GAGGTGGAGG CCGCGATGGC GGCCTACCGC ACTGATGTCC	300
CCATCACACC GTGCGAACTC ACGGCGGCTA AAAAGGCCGC CCAACAGCTG GTATTGTCCG	360
CGACAACAT GCGGGAATAC CTGGCGGCCG GTGCCAAGA GCGGCAGCTT CTGCGACCTT	420
CCCTGGCCAA CCGGCCCAAG GCCTATGCCG AGCTTGATCA GGAGGCTCGC ACCGCGCTGC	480
ACAACGACGG CGAAGCAACT GTGCAGGCAG AATCGGCCGG GCGCGTCGGA GGGGACACTT	540
CGGCCCACT AACCGATACG CGAGGGGTGG CCACGGGCCG TGAAGCCAAC TTGATGGATC	600
TCAAGAGAGC GGCAAGCAAG CTCGAAACCG GCGACCAAGC CGCATCGCTC GCGCACTTTG	660
CGGATGGGTG GAACACTTTC AACCTGACGC TCGAAGGCGA CCGCAAGCGG TTCGCGGGCT	720
TTGACAACCG GGAAGGGGAT GCGGCTACCG CTTCGGAGGC TTGCTCGAT CAACAACGGC	780
AATGGATACT CCACATGGCC AATGTAGCGC CTGCGATGGC CAAGCAGGCT CAATATGTCC	840
CGCAGCTGCA CGTGTGGGCT AGCGGGCAAC ATCCGACTTA TGAAGACATA GTCGGCTCG	900
AACGGCTTTA CGCGGAAAC CTTTCGGCCC GCGACCAAAAT TCTCCCGGTG TACCGCGAGT	960
ATCAGCAGAG GTGGGAGAAG GTGCTGACCG AATACAACAA CAAGGCAGCC CTGGAACCGG	1020
TAAACCCGCC GAAGCCTCCC CCGGCAATCA AGATCGACCC GCGCCCGGCT CCGCAAGAGC	1080
AGGGATTGAT CCTTGGCTTC CTGATGCCGC CGTCTGACGG CTCGGGTGTG ACTCCCGGTA	1140

```

CCGGGATGCC ACCCGCACCG ATGGTTCCGC CTACCGGATC GCGGGGTGGT GGCTTCCCGS 1200
CTGACACGGC GCGGCAGCTG ACGTCCGCTE GGCGGGAAGC CUCAGCGGTG TCGGGCGACG 1250
TGGCGGTCAA ACCGSCATCG CTCGCTGCCG GTGGAGSCCG GGGGTTGCCS TCGCGCCCGT 1320
TGGGATCCUC GATCGGGGSC GCGGAATCGG TCGGSCCCGC TGGGCTGGT GACATTGCCG 1380
GCTTAGGCCA GGAAGGGCC GCGGCGGCG CCGCGCTGGC GGGCGGTGGC ATGGGAATGC 1440
CGATGGGTGC CCGGCATCAG GCACAAGGGG GCGCCAGTC CAAGGTTCT CAGCAGGAAG 1500
ACGAGGCGCT CTACACCGAG GATCGGGCAT GGACCGAGGC CGTCATTGGT AACCGTCGGC 1560
GCCAGGACAG TAGGASTCG AAGTAGCAT GGACGATTC GACCGGCTG TCGCCCGGCT 1620
GTTGACGCTG CCGCGCGCGT TTCAGTCGGC CCTAGACGGG ACGCTCAATC AGATGAACAA 1680
CGGATCTTTC CCGCGCCAGC ACGAAGCCGA GACGCTCGAA GTCAGGATCA ATGGGCACCA 1740
GTGGCTCACC GGCTTGGCA TCGAAGATGG TTGCTGAAG AAGCTGGTG CCGAGGCGGT 1800
GGCTCAGCGG GTCAACGAGG CGCTGCACAA TGCCGAGGCC GCGGCTCCG CGTATAACGA 1860
CGCGGCGGCG GAGCAGCTGA CCGCTGCGTT ATCGGCCATG TCCCGCGCGA TGAACGAAGC 1920
AATGGCCYAA GCGCATTTST GCGTGGTAG GGAATACGA CGAATGAGC GCGGCAATGC 1980
GGTCATTCAE CCGCGCCGAC ACGGCGTGAG TACCATTTGT CAATGTTTG ACATGGATCG 2040
GCGGGTTCC GAGGGGCGCA TAGTCTGGT CCGCAATATT GCGGAGCTA GCTGGTCTTA 2100
GGTTCGGTTA CCGTGGTTAA TTATGACGTC CGTTACCA 2138

```

## (2) INFORMATION FOR SEQ ID NO:179:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 460 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:179:

```

Met Thr Gln Ser Gln Thr Val Thr Val Asp Gln Gln Glu Ile Leu Asn
1           5           10           15

Arg Ala Asn Glu Val Glu Ala Pro Met Ala Asp Pro Pro Thr Asp Val
20           25           30

Pro Ile Thr Pro Cys Glu Leu Thr Ala Ala Lys Asn Ala Ala Gln Gln

```

190

35	40	45	
Leu Val Leu Ser Ala Asp Asn Met Arg Glu Tyr Leu Ala Ala Gly Ala			
50	55	60	
Lys Glu Arg Gln Arg Leu Ala Thr Ser Leu Arg Asn Ala Ala Lys Ala			
65	70	75	80
Tyr Gly Glu Val Asp Glu Glu Ala Ala Thr Ala Leu Asp Asn Asp Gly			
	85	90	95
Glu Gly Thr Val Gln Ala Glu Ser Ala Gly Ala Val Gly Gly Asp Ser			
	100	105	110
Ser Ala Glu Leu Thr Asp Thr Pro Arg Val Ala Thr Ala Gly Glu Pro			
	115	120	125
Asn Phe Met Asp Leu Lys Glu Ala Ala Arg Lys Leu Glu Thr Gly Asp			
	130	135	140
Gln Gly Ala Ser Leu Ala His Phe Ala Asp Gly Trp Asn Thr Phe Asn			
	145	150	155
Leu Thr Leu Gln Gly Asp Val Lys Arg Phe Arg Gly Phe Asp Asn Trp			
	165	170	175
Glu Gly Asp Ala Ala Thr Ala Cys Glu Ala Ser Leu Asp Gln Glu Arg			
	180	185	190
Gln Trp Ile Leu His Met Ala Lys Leu Ser Ala Ala Met Ala Lys Gln			
	195	200	205
Ala Gln Tyr Val Ala Gln Leu His Val Trp Ala Arg Arg Glu His Pro			
	210	215	220
Thr Tyr Glu Asp Ile Val Gly Leu Glu Arg Leu Tyr Ala Glu Asn Pro			
	225	230	235
Ser Ala Arg Asp Gln Ile Leu Pro Val Tyr Ala Glu Tyr Gln Gln Arg			
	245	250	255
Ser Glu Lys Val Leu Thr Glu Tyr Asn Asn Lys Ala Ala Leu Glu Pro			
	260	265	270
Val Asn Pro Pro Lys Pro Pro Pro Ala Ile Lys Ile Asp Pro Pro Pro			
	275	280	285
Pro Pro Glu Glu Gln Gly Leu Ile Pro Gly Phe Leu Met Pro Pro Ser			
	290	295	300
Asp Gly Ser Gly Val Thr Pro Gly Thr Gly Met Pro Ala Ala Pro Met			
	305	310	315
Val Pro Pro Thr Gly Ser Pro Gly Gly Gly Leu Pro Ala Asp Thr Ala			
	325	330	335

```

Ala Gln Leu Thr Ser Ala Gly Arg Glu Ala Ala Ala Leu Ser Gly Asp
    340                      345                      350

Val Ala Val Lys Ala Ala Ser Leu Gly Gly Gly Gly Gly Gly Val
    355                      360                      365

Pro Ser Ala Pro Leu Gly Ser Ala Ile Gly Gly Ala Glu Ser Val Arg
    370                      375                      380

Pro Ala Gly Ala Gly Asp Ile Ala Gly Leu Gly Gln Gly Arg Ala Gly
    385                      390                      395                      400

Gly Gly Ala Ala Leu Gly Gly Gly Gly Met Gly Met Pro Met Gly Ala
    405                      410                      415

Ala His Gln Gly Gln Gly Gly Ala Lys Ser Lys Gly Ser Gln Gln Glu
    420                      425                      430

Asp Glu Ala Leu Tyr Thr Glu Asp Arg Ala Trp Thr Glu Ala Val Ile
    435                      440                      445

Gly Asn Arg Arg Arg Gln Asp Ser Lys Glu Ser Lys
    450                      455                      460

```

## (2) INFORMATION FOR SEQ ID NO:180:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 277 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:180:

```

Ala Gly Asn Val Thr Ser Ala Ser Gly Pro His Arg Phe Gly Ala Pro
1           5           10           15

Asp Arg Gly Ser Gln Arg Arg Arg Arg His Pro Ala Ala Ser Thr Ala
20           25           30

Thr Glu Arg Cys Arg Phe Asp Arg His Val Ala Arg Gln Arg Cys Gly
35           40           45

Phe Pro Pro Ser Arg Arg Gln Leu Arg Arg Arg Val Ser Arg Glu Ala
50           55           60

Thr Thr Arg Arg Ser Gly Arg Arg Asn His Arg Cys Gly Trp His Pro
65           70           75           80

Gly Thr Gly Ser His Thr Gly Ala Val Arg Arg Arg His Gln Gln Ala

```



192

				85				90				95			
Arg	Asp	Gln	Ser	Leu	Leu	Leu	Arg	Arg	Arg	Gly	Arg	Val	Asp	Leu	Asp
100				105				110							
Gly	Gly	Gly	Arg	Leu	Arg	Arg	Val	Tyr	Arg	Phe	Gln	Gly	Cys	Leu	Val
115				120				125							
Val	Val	Phe	Gly	Gln	His	Leu	Leu	Arg	Pro	Leu	Leu	Ile	Leu	Arg	Val
130				135				140							
His	Arg	Glu	Asn	Leu	Val	Ala	Gly	Arg	Arg	Val	Phe	Arg	Val	Lys	Pro
145				150				155				160			
Phe	Glu	Pro	Asp	Tyr	Val	Phe	Ile	Ser	Arg	Met	Phe	Pro	Pro	Ser	Pro
165				170				175							
His	Val	Gln	Leu	Arg	Asp	Ile	Leu	Ser	Leu	Leu	Gly	His	Arg	Ser	Ala
180				185				190							
Gln	Phe	Gly	His	Val	Glu	Tyr	Pro	Leu	Pro	Leu	Leu	Ile	Glu	Arg	Ser
195				200				205							
Leu	Ala	Ser	Gly	Ser	Arg	Ile	Ala	Phe	Pro	Val	Val	Lys	Pro	Pro	Glu
210				215				220							
Pro	Leu	Asp	Val	Ala	Leu	Gln	Arg	Glu	Val	Glu	Ser	Val	Pro	Pro	Ile
225				230				235				240			
Arg	Lys	Val	Arg	Glu	Arg	Cys	Ala	Leu	Val	Ala	Arg	Phe	Glu	Leu	Pro
245				250				255							
Cys	Arg	Phe	Phe	Glu	Ile	His	Glu	Val	Gly	Phe	Thr	Gly	Arg	Gly	His
260				265				270							
Pro	Arg	Arg	Ile	Gly											
275															

## (2) INFORMATION FOR SEO IS NO. 101:

## 11) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 192 amino acids  
(B) TYPE: amino acid  
(C) STRANGENESS:  
(D) TOPOLOGY: linear

{x1} SEQUENCE DESCRIPTION: SEQ ID NO:181:

Arg Val Ala Ala Ser Phe Ile Asp Trp Leu Asp Ser Pro Asp Ser Pro  
1                   5                   10                   15

193

```

Leu Asp Pro Ser Leu Val Ser Ser Leu Leu Asn Ala Val Ser Cys Gly
    20                      25                      30

Ala Glu Ser Ser Ala Ser Ser Ser Ala Arg Ser Gly Asn Gly Ser Arg
    35                      40                      45

Trp Thr Ser Met Pro Ser Gly Thr Arg Pro Gly Pro Arg Arg Ala Thr
    50                      55                      60

Ser Arg Asp Asp Arg Arg Ser Ala Thr Ser Val Ile Pro Ser Arg Arg
    65                      70                      75                      80

Ser Val Ala Pro Arg Ala Glu Phe Gly Thr Arg Leu Ala Ser His Arg
    85                      90                      95

Ala Ser Pro Ser Asn Ala Cys Pro Val Arg Ile Val Thr Ser Ala Ser
   100                      105                      110

Gly Arg Pro Ile Ser Ser Pro Pro Ile Val Arg Ser Arg Ser Cys Val
   115                      120                      125

Asp Lys Asn Gly Arg Arg Cys Ala Ser Gly Tyr Arg Arg Leu Asn Arg
   130                      135                      140

Ala Arg Ser Ser Ser Ile Ala Ala Arg Cys Arg Thr Ile Gly Thr Phe
   145                      150                      155                      160

Arg Arg Ser Arg Tyr Ser Ala Ser Met Arg Val Ser Thr Asn Ser Pro
   165                      170                      175

His Val Thr His Gly Val Ala Pro Gly Val Thr Arg Arg Ile Gly Gly
   180                      185                      190

```

## (2) INFORMATION FOR SEQ ID NO:182:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 196 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:182:

```

Gln Gln Arg Pro Gln Met Cys Gln Arg Val Ser Glu Ile Glu Pro Arg
1          5          10          15

Thr Gln Phe Phe Asn Arg Cys Ala Leu Pro His Tyr Trp His Phe Pro
    20          25          30

Ala Val Ala Val Phe Ser Lys His Ala Ser Leu Asp Glu Leu Ala Pro

```

	35		40		45	
Arg	Asn	Pro	Arg	Arg	Ser	Ser
50					55	
Ile	Phe	Ala	Ala	Thr	Leu	Val
65				70		
Gly	Gly	Glu	Ala	Asp	Gln	Leu
			85			90
Ala	Gly	Arg	Val	Arg	Arg	Gly
			100			105
Glu	Gly	Ala	Ala	Pro	Asp	Leu
			115			120
Gln	Gln	His	Arg	Arg	Leu	Gln
			130			135
Glu	Thr	Ala	His	His	Arg	Arg
			145			150
His	Val	Arg	Gly	Pro	Ser	Asn
			165			170
Arg	His	Ser	Gly	Ala	Gly	Gly
			180			185
Gly	Gly	Ser	Ala			
			195			

## (2) INFORMATION FOR SEQ ID NO:183:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 311 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (vi) SEQUENCE DESCRIPTION: SEQ ID NO:183:

Val	Arg	Cys	Gly	Thr	Leu	Val	Pro	Val	Pro	Met	Val	Glu	Phe	Leu	Thr
1				5					10					15	
Ser	Thr	Asn	Ala	Pro	Ser	Leu	Pro	Ser	Ala	Tyr	Ala	Glu	Val	Asp	Lys
			20					25					30		
Leu	Ile	Gly	Leu	Pro	Ala	Gly	Thr	Ala	Lys	Arg	Trp	Ile	Asn	Gly	Tyr
		35					40					45			

195

Glu	Arg	Gly	Gly	Lys	Asp	His	Pro	Pro	Ile	Leu	Arg	Val	Thr	Pro	Gly	50	55	60
Ala	Thr	Pro	Trp	Val	Thr	Trp	Gly	Glu	Phe	Val	Glu	Thr	Arg	Met	Leu	65	70	75
Ala	Glu	Tyr	Arg	Asp	Arg	Arg	Lys	Val	Pro	Ile	Val	Arg	Glu	Arg	Ala	85	90	95
Ala	Ile	Glu	Glu	Leu	Arg	Ala	Arg	Phe	Asn	Leu	Arg	Tyr	Pro	Leu	Ala	100	105	110
His	Leu	Arg	Pro	Phe	Leu	Ser	Thr	His	Glu	Arg	Asp	Leu	Thr	Met	Gly	115	120	125
Gly	Glu	Glu	Ile	Gly	Leu	Pro	Asp	Ala	Glu	Val	Thr	Ile	Arg	Thr	Gly	130	135	140
Gln	Ala	Leu	Leu	Gly	Asp	Ala	Arg	Trp	Leu	Ala	Ser	Leu	Val	Pro	Asn	145	150	155
Ser	Ala	Arg	Gly	Ala	Thr	Leu	Arg	Arg	Leu	Gly	Ile	Thr	Asp	Val	Ala	165	170	175
Asp	Leu	Arg	Ser	Ser	Arg	Glu	Val	Ala	Arg	Arg	Gly	Pro	Gly	Arg	Val	180	185	190
Pro	Asp	Gly	Ile	Asp	Val	His	Leu	Leu	Pro	Phe	Pro	Asp	Leu	Ala	Asp	195	200	205
Asp	Asp	Ala	Asp	Asp	Ser	Ala	Pro	His	Glu	Thr	Ala	Phe	Lys	Arg	Leu	210	215	220
Leu	Thr	Asn	Asp	Gly	Ser	Asn	Gly	Glu	Ser	Gly	Glu	Ser	Ser	Gln	Ser	225	230	235
Ile	Asp	Asp	Ala	Ala	Thr	Arg	Tyr	Met	Thr	Asp	Glu	Tyr	Arg	Gln	Phe	245	250	255
Pro	Thr	Arg	Asn	Gly	Ala	Gln	Arg	Ala	Leu	His	Arg	Val	Val	Thr	Leu	260	265	270
Leu	Ala	Ala	Gly	Arg	Pro	Val	Leu	Thr	His	Cys	Phe	Ala	Gly	Lys	Asp	275	280	285
Arg	Thr	Gly	Phe	Val	Val	Ala	Leu	Val	Leu	Glu	Ala	Val	Gly	Leu	Asp	290	295	300
Arg	Asp	Val	Ile	Val	Ala	Asp										305	310	

(2) INFORMATION FOR SEQ ID NO:184:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2672 base pairs

(B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:184:

CTCTGCCCCA TTCCGCACGA GCTGAGCAGC CCAAGGGGCC GTTCGGCBA A GTCATCGAGG	60
CATTCGCCGA CGGCTGEEC GGCAGGETA AGCAATCAA CACCACGCTG AACAGCCTGT	120
CGCAGGCGTT GRACGCTTG AATGAGGGCC GCGCGACTT CTTCGCGGTG GTACCCAGCC	180
TGGCGCTATT CGTCAACGGG CTACATCAGG AGGACCAACA GTTCGTGGCG TTGAACAAGA	240
ACCTTGGCGA GTTCAAGGAC AGGTTGACCC ACTCCGATCC GGACCTGTCC AACGCCATCC	300
AGCAATTCGA CAGCTTGGTC GCCGTGCGCC GCGGTTCTT GCGCAAGAAC CGCGAGGTCC	360
TGACCCATGA CGTCAATAT CTGCGGACCG TGACCACCCAC GTTCTGTGAG CCGGATCCCT	420
TGGATGGGTT GGAGACCGTC CTGCACATCT TCCCGAGCGT GCGCGCGAAC ATTAACCAGC	480
TTTACCATCC GACACAGGCT GCGGTGGTGT CGCTTTCCGC GTTCACGAAT TTCGCGACCC	540
CGATGGAGTT CATCTGCAGC TCGATTCAGG CGGCTAGCCG GCTCGGTTAT CAGAGTCCG	600
CGGAATCTG TCGGAGTAT CTGGCGCCAG TCCTCGATGC GATCAAGTTC AACTACTTTC	660
CGTTCCGGCT GAACGTGGCC AGCAGCGGCT CGACACTGCC TAAAGAGATC GCGTACTCCG	720
AGCCCCGCTT GCAGCGGCC AACGGGTACA AGGACCCAC GGTGCCCCGC ATCTGGGTCC	780
CGGATACGCC GTTGTACAC CGCACACCC AGCCCGGTTG GGTGGTGGCA CCGGGGATCC	840
AAGGGGTTCA GGTGGGACCG ATCAGCGAGG GTTGTCTGAC GCGGAGTCC CTGGCCGAC	900
TCATGGGTGG TCCGATATC GCGGCTCGCT COTCAGGGCT GCAAGCCCCG CCGGACCCC	960
CGAATGGETA CGACGAGTAC CCGTGTCTGC CCGGATCCG TTTACAGGCC CCACAGGTCC	1020
CGATACCACC GCGGCTTCTT GGGCCGAGC TAATCCCGG TCGGTTGCCA CCGGTCTTGG	1080
CGGCGATCGT GTTCCCAAGA GATCGCCCG CAGGTCGGA AACTTCCGAC TACATGGGTC	1140
TCTGTCTGCT GTCCCGGGCC CTGGGAGCCT TCTGTTCGG GGTGTCTCT AGCCCCGGCC	1200
GTGGAAGGAT GGCGATCCG CAGGTGTGA TACCGCGGAT CACCGGCTG GCGTTGATCC	1260
CGGCATTCTT CGCACATTG TGGTACCGCA CAGAACATCC GTCATAGAC ATGCGCTTGT	1320
TCCAGAACCG AGCGGTGGG CAGGCCAACA TGACGATGAC GGTGCTTCTC CTCGGGCTGT	1380

```

TTGGCTCCTT CTTGCTGCTC CCGAGCTACC TCCAGCAAGT GTTGCACCAA TCACCGATGC 1440
AATCGGGGGT GCATATCATC CCACAGGGCC TCGGTGCCAT GGTGGCGATG CCGATCGCCG 1500
GAGCGATGAT GGACCGACCG GGACCGGCCA AGATCGTGCT GGTGGGATC ATGCTGATCG 1560
CTCGGGGGTT GGCGACCTTC GCGTTTGGTG TCGCGCGCCA AGCGGACTAC TTACCCATTC 1620
TCCCGACCCG GCTGGCAATC ATCGGCATGG GCATGGGCTG CTCCATGATG CCACTGTCCG 1680
GGGCGGCAGT GCAGACCTG GCGGCACATC AGATCGCTCG CGGTTCGACG CTGATCAGCG 1740
TCAACCAGCA GGTGGCGGT TCGATAGGGA CCGCACTGAT GTCGGTGCTG CTCACCTACC 1800
AGTTCAATCA CAGCGAAATC ATCGCTACTG CAAAGAAAGT CGCACTGACC CCAGAGAGTG 1860
GCGCGGGGCG GGGGGGCGCG GTTGACCTT CCTCGCTACC GCGCGAAGC AACTTCGCGG 1920
CCCACTGCT GCATGACCT TCGCACGCT AGCGGCTGGT ATTGCTGATA GCGACCGCGC 1980
TAGTGGTCTC GACGCTGAT CCGCGGCGAT TCTGCGGAA ACAGCAGCT AGTCATCGAA 2040
GAGCACCGTT GCTATCGCA TGACGTCTTC TT 2072

```

(2) INFORMATION FOR SEQ ID NO:185:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1923 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:185:

```

TCACCCCGGA GAAGTCGTTG GTCACGACC TGGACATCGA CTGGCTGTCTG ATGCTCGAGA 60
TGGCGTGCA GACCGAGGAC AAGTACGGCG TCAAGATCCC CGACGAGGAC CTCGCGGCTC 120
TGGTACCGT CGGTGAGGTT GTCGCTACA TCCGGAAGCT CGACGAGAA AACCCGAGG 180
CGGCTCAGGC GTTGGCGCGG AGATTTGAGT CCGAGAACCC CGATCGCGCA CGAGCAGATC 240
GGTGGGTTTC ACCACATCG CAAGCTGAG AGCGCGTCC TCTCTTGCA CGCTCAGCCA 300
GGTGGCGTG TCGCGGCTT CCAGCAAGTG TTCCACCCAC ACGAGGGGAC CCTCGCGAAA 360
GGTCACTGAT CTGCGACCA CATAGTCGAT GCGACGCTG CTGACATTC GCGCGGCTCC 420
GAGTTGGCGG GGGCCGAATT GCGCAATTC GTCGAGGCC AGCGGATCCC GCGCGCGCGC 480

```

CGBGCTGGCT	GGTGTTTTGG	GCCGCCGGAT	GCCCACGACG	AGAACGACGA	TGGCGGGCAT	540
GAACAGCGCC	ACGGCAATCA	CGACACGACG	ATTTCGCACG	CATACCTCTT	CGTACCGCTG	600
CGCGCGGCTT	GGTCGATCGG	TGCATATCGG	ATGCGCGCCT	TTAAGGTAAC	AGCTTTGCGG	660
GGATCGGGGG	TCACAAAGGG	CGACTTGTCC	GGCGCGGAAC	CCGCGAGGTC	TGGCGCGCGG	720
TCACCCGAGC	TCACTGGTGC	ACCATCGCGG	TGTGGGTGAG	CGTGCNACTC	AAACACACTC	780
AACGGCAACG	GTTCCTCAGG	TCACGAGGTC	AACCTCGACC	CGCAATCGCT	CGTACGTTTC	840
GACCGCGCGC	AGGTGCGGAG	TCAGCAGGTT	TGCGCGGCGA	GCTTTCGCGG	TGAAGCGGAC	900
CAGGGGCTCG	TAGGTTGCGC	CACCGGTGAC	ATCGTGCTCG	GCGAGGTGGT	CGGTCAAGGC	960
GCGATATGAG	CAGGCATCCA	GTGCCAGGTA	GTTCCTGGAG	GTGATGTCCG	CGAAGTAGGC	1020
GTGGACGGCA	ACAGGGGCAA	TACGATGCGG	CGGTGGTAAC	CGGCTCAGAG	CCGAATAGGT	1080
TTCCACAGCC	GCGTGGCGGA	TCAGATGGAC	GCCACGGTTG	AGCGCGCGCA	CGGCGGCTTC	1140
GTGCGCTTCS	TGCCAGGTCG	CGAATCGGCG	AACGAGGACG	CTGGTGTCTG	GTGCGATCAC	1200
CGGCTGTGTC	GATCGAGGCT	TTCCCGAAGG	ATTTCGTGCG	TCAACGGGCG	CAGGGGAGCT	1260
TCGCGCGGTG	CGACGAGAAC	CGAGGCTTCC	CGAAGCAGTT	CGACACCGGT	CGGGGCGGCG	1320
TCATCTCGA	TGCGGCGATC	GCGCTCGGTC	ATCTCCACCT	GGTGGTTCCC	GCGCAGGCGA	1380
AGGCGCTCGC	GAATCGGCTT	GCGAATCACC	AGAGCTCCTG	CGACATCGAT	GGTTGTTCGC	1440
ATGGTAGGAA	ATTACCAATC	GCACGTTCCA	TAGGCGGTGC	CTGGCGGCGA	TGTGCGGACG	1500
ATCGGCTAGC	GATCGGAACG	ATTGTTTCGG	AAATGGCTGA	GGGAGCGTGC	GGTGCGGGTC	1560
ATGGGTCTCG	ATCCCGGCTT	GACCGGATGC	GCGCTGTGCG	TCATCGAGAG	TGGGCGTGGT	1620
CGGCAGCTCA	CGCGGCTGGA	TGTGAGGTCG	GTGCGCACAC	CGTGGGATGC	GGGCTTGGCG	1680
CAGCGGCTGT	TGGCCATCAG	CGATGCGGTC	GAGCACTGGC	TGGACACCCA	TCATCGGAGG	1740
GTGGTGCTTA	TGGAACGGGT	GTTCCTTCAG	CTCAACGTGA	CCACGGTGAAT	GGGCAAGGCG	1800
CAGGCGGGCG	GCGTGAATCG	CCTGCGGCGG	GCGAAGGTCG	GTGTGCAAGT	GAATTTCCAT	1860
ACCGCCAGCG	AGTCAAGGCG	GCGGTCACCT	GCGAAGGCTT	CGGCAGACAA	GCGTCAGGTC	1920
ACC						1923

{2} INFORMATION FOR SEQ ID NO:186:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1055 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:186:

CTGGCGTGGC AGTGTACCG GCGATATGAC GTCGGCATTC AATTCGGCGG CCCCAGCGGA	60
CCCGTGGCCA CCAATCTGG ACCACCCGGT CCGTCAATTG CCGAAGGTGG CCAAGTGGGT	120
GCCCAATGTG GTGCTGGGTT TCTTGAACGA AGGCTTGGCG TATCGGGTGC CCTACCCCCA	180
AACAACGCCA GTCCAGGAAT CCGGTCCCGC GCGGCCGATT CCCAGCGGCA TGTGCTAGCC	240
GGGGATGGTT CAGACGTAACT GGTGGCTAG GTCGAAACCC GCGCCAGGAC CCTGGAGCG	300
GCTCATGSCA GCGAAATTAG AAAACCGGGG ATATTGTCCG CGGATTGTCA TACGATGCTG	360
AGTGCTTGGT GGTTCGTGTT TGGCATTTGA GTGTGATGTT GTTGAGACCC TGGCCTGGAA	420
GGGGACAAAG TGCTTTTGGC TCTTGGTCCG CCTTTGCCGC CCGACGCGGT GGTGGCGAAA	480
CGGGCTGAGT CCGGAATGCT CCGCGGGTTC TCGGTTCCGC TCAGCTGGGG AGTGGCTGTG	540
CCACCCGATG ATTATGACCA CTGGGCGGCT GCGGCGGAGG ACGGCGCGGA TGTGATGTC	600
CAGGCGGCGG AAGGGCGGGA CCGAGAGGCG GCGGCCATGC ACGAGTGGGA TGAGTGGCAG	660
GCTGGAACG AGTGGGTGGC GGAGAACGCT GAAACCCGCT TTGAGGTGGC ACGGACTAGC	720
AGCAGCGTGA TTCCGCATTC TCCGCGCGCC GGTAGGAGA GGGGCGCGAG ACTGTCTTA	780
TTTGACCACT GATCGGCGGT CTCGGTGTTC CCGCGGCGCG CTATGACAAC AGTCATGTG	840
CATGACCACT TACAGGTTT AGGTCCAGGT TCAACAGGCA GACAGGCAAC ATGSCAACAC	900
GTTTTATGAC GGATCCGCAC GCGATCGGGG ACATGCGCGG CCGTTTTGAG GTGCACGCCC	960
AGAGGTTGGA GGACGAGGCT CCGCGGATGT GGGGCTCCCG GCAAAACATC TCGGCGCGCG	1020
GCTGGAGTGG CATGCGCGAG GCGACCTCGC TAGAC	1055

## (2) INFORMATION FOR SEQ ID NO:187:

## (1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 359 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear



## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:187:

CCGCGCTCCTT	CTTGGCATAC	TCCGCGCGCG	CCSCCTCGAC	CGCACTGGCC	GTGCGCTGTG	60
TCCGGGCTGA	CCACUGGGAT	CGCCGAACCA	TCCGAGATCA	CCTCGCAATG	ATCCACCTCG	120
CGCAGCTGGT	CACCCAGCCA	CCGGGCGGTG	TGCGACAGCG	CCTGCATCAC	CTTGGTATAG	180
CGTCGGGCCC	CCAGCCGCAG	GAGCTTGTAG	TACTGGCCCC	CCACCTGGTY	ACCGGGACGG	240
GAGAAGTICA	GGGTGAAGGT	CGGCATGTGG	CGCCCGAGGT	AGTTGACCGG	GAAAACCCAG	300
TCTCCGGGCA	GGTCTCGGG	CCGCGCCAC	ACGACAAACC	CGACGCGCGG	ATAGGTCAG	359

## (2) INFORMATION FOR SEQ ID NO:188:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 350 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:188:

AACGGGCCCC	TGGGCACGCG	TCTCTAAGG	GCTCTCGTTG	GTGGCATGAA	GTGCTGGAAG	60
GATGCTCTCT	GGCAGATTCC	CGCCAGAGCA	AAACAGCCCG	TAGTCCTAGT	CCGAGTCGCC	120
CGCAAGCTTC	CTCGAATAAC	TCCGTACCCG	GAGCGCCAAA	CCGGGTCTCC	TTGGCTAAGC	180
TGCGCGAACC	ACTTGAGGTY	CCCGACTCC	TTGACGTCCA	GACCGATTCC	TTCGAGTGGC	240
TGATCGGTTC	GCCGCGCTGG	CGCGAATCCG	CCGCGGAGCG	GGGTGATGTC	AACCCAGTGG	300
GTGGCCTGGA	AGAGGTGCTC	TACGAGCTGT	CTCCGATCCA	CGACTCTGCC		350

## (2) INFORMATION FOR SEQ ID NO:189:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 579 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:189:

201

Glu	Gln	Pro	Lys	Gly	Pro	Phe	Gly	Gln	Val	Ile	Gln	Ala	Phe	Ala	Asp		
1				5					10					15			
Gly	Leu	Ala	Gly	Lys	Gly	Lys	Gln	Ile	Asn	Thr	Thr	Leu	Asn	Ser	Leu		
			20					25					30				
Ser	Gln	Ala	Leu	Asn	Ala	Leu	Asn	Gln	Gly	Arg	Gly	Asp	Phe	Phe	Ala		
		35					40					45					
Val	Val	Arg	Ser	Leu	Ala	Leu	Phe	Val	Asn	Ala	Leu	His	Gln	Asp	Asp		
	50					55					60						
Gln	Gln	Phe	Val	Ala	Leu	Asn	Lys	Asn	Leu	Ala	Gln	Phe	Thr	Asp	Arg		
65				70						75					80		
Leu	Thr	His	Ser	Asp	Ala	Asp	Leu	Ser	Asn	Ala	Ile	Gln	Gln	Phe	Asp		
			85						90					95			
Ser	Leu	Leu	Ala	Val	Ala	Arg	Pro	Phe	Phe	Ala	Lys	Asn	Arg	Gln	Val		
			100					105					110				
Leu	Thr	His	Asp	Val	Asn	Asn	Leu	Ala	Thr	Val	Thr	Thr	Thr	Leu	Leu		
	115					120					125						
Gln	Pro	Asp	Pro	Leu	Asp	Gly	Leu	Gln	Thr	Val	Leu	His	Ile	Phe	Pro		
	130					135					140						
Thr	Leu	Ala	Ala	Asn	Ile	Asn	Gln	Leu	Tyr	His	Pro	Thr	His	Gly	Gly		
145				150					155					160			
Val	Val	Ser	Leu	Ser	Ala	Phe	Thr	Asn	Phe	Ala	Asn	Pro	Met	Gln	Phe		
			165					170					175				
Ile	Cys	Ser	Ser	Ile	Gln	Ala	Gly	Ser	Arg	Leu	Gly	Tyr	Gln	Gln	Ser		
		180					185					190					
Ala	Gln	Leu	Cys	Ala	Gln	Tyr	Leu	Ala	Pro	Val	Leu	Asp	Ala	Ile	Lys		
	195					200					205						
Phe	Asn	Tyr	Phe	Pro	Phe	Gly	Leu	Asn	Val	Ala	Ser	Thr	Ala	Ser	Thr		
	210					215					220						
Leu	Pro	Lys	Gln	Ile	Ala	Tyr	Ser	Gln	Pro	Arg	Leu	Gln	Pro	Pro	Asn		
225			230						235					240			
Gly	Tyr	Lys	Asp	Thr	Thr	Val	Pro	Gly	Ile	Trp	Val	Pro	Asp	Thr	Pro		
			245					250					255				
Leu	Ser	His	Arg	Asn	Thr	Gln	Pro	Gly	Trp	Val	Val	Ala	Pro	Gly	Met		
		260						265				270					
Gln	Gly	Val	Gln	Val	Gly	Pro	Ile	Thr	Gln	Gly	Leu	Leu	Thr	Pro	Gln		
	275					280					285						

202

Ser Leu Ala Glu Leu Met Gly Gly Pro Asp Ile Ala Pro Pro Ser Ser  
 280 290 300  
 Gly Leu Gln Thr Pro Pro Gly Pro Pro Asn Ala Tyr Asp Glu Tyr Pro  
 305 310 315 320  
 Val Leu Pro Pro Ile Gly Leu Gln Ala Pro Gln Val Pro Ile Pro Pro  
 325 330 335  
 Pro Pro Pro Gly Pro Asp Val Ile Pro Gly Pro Val Pro Pro Val Leu  
 340 345 350  
 Ala Ala Ile Val Phe Pro Arg Asp Arg Pro Ala Ala Ser Glu Asn Phe  
 355 360 365  
 Asp Tyr Met Gly Leu Leu Leu Leu Ser Pro Gly Leu Ala Thr Phe Leu  
 370 375 380  
 Phe Gly Val Ser Ser Ser Pro Ala Arg Gly Thr Met Ala Asp Arg His  
 385 390 395 400  
 Val Leu Ile Pro Ala Ile Thr Gly Leu Ala Leu Ile Ala Ala Phe Val  
 405 410 415  
 Ala His Ser Trp Tyr Arg Thr Glu His Pro Leu Ile Asp Met Arg Leu  
 420 425 430  
 Phe Gln Asn Arg Ala Val Ala Gln Ala Asn Met Thr Met Thr Val Leu  
 435 440 445  
 Ser Leu Gly Leu Phe Gly Ser Phe Leu Leu Leu Pro Ser Tyr Leu Gln  
 450 455 460  
 Gln Val Leu His Gln Ser Pro Met Gln Ser Gly Val His Ile Ile Pro  
 465 470 475 480  
 Gln Gly Leu Gly Ala Met Leu Ala Met Pro Ile Ala Gly Ala Met Met  
 485 490 495  
 Asp Arg Arg Gly Pro Ala Lys Ile Val Leu Val Gly Ile Met Leu Ile  
 500 505 510  
 Ala Ala Gly Leu Gly Thr Phe Ala Phe Gly Val Ala Arg Gln Ala Asp  
 515 520 525  
 Tyr Leu Pro Ile Leu Pro Thr Gly Leu Ala Ile Met Gly Met Gly Met  
 530 535 540  
 Gly Cys Ser Met Met Pro Leu Ser Gly Ala Ala Val Gln Thr Leu Ala  
 545 550 555 560  
 Pro His Gln Ile Ala Arg Gly Ser Thr Leu Ile Ser Val Asn Gln Gln  
 565 570 575  
 Val Gly Gly Ser Ile Gly Thr Ala Leu Met Ser Val Leu Leu Thr Tyr

203

```

          580          585          590
Gln Phe Asn His Ser Glu Ile Ile Ala Thr Ala Lys Lys Val Ala Leu
    595          600          605

Thr Pro Glu Ser Gly Ala Gly Arg Gly Ala Ala Val Asp Pro Ser Ser
    610          615          620

Leu Pro Arg Gln Thr Asn Phe Ala Ala Gln Leu Leu His Asp Leu Ser
    625          630          635          640

His Ala Tyr Ala Val Val Phe Val Ile Ala Thr Ala Leu Val Val Ser
    645          650          655

Thr Leu Ile Pro Ala Ala Phe Leu Pro Lys Gln Gln Ala Ser His Arg
    660          665          670

Arg Ala Pro Leu Leu Ser Ala
    675

```

## (2) INFORMATION FOR SEQ ID NO:199:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 120 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:199:

```

Thr Pro Glu Lys Ser Phe Val Asp Asp Leu Asp Ile Asp Ser Leu Ser
1          5          10          15

Met Val Glu Ile Ala Val Gln Thr Glu Asp Lys Tyr Gly Val Lys Ile
    20          25          30

Pro Asp Glu Asp Leu Ala Gly Leu Arg Thr Val Gly Asp Val Val Ala
    35          40          45

Tyr Ile Gln Lys Leu Glu Glu Gln Asn Pro Gln Ala Ala Gln Ala Leu
    50          55          60

Arg Ala Lys Ile Glu Ser Glu Asn Pro Asp Ala Ala Arg Ala Asp Arg
    65          70          75          80

Cys Val Ser Pro Thr Ser Gln Ala Arg Asp Ala Arg Arg Pro Leu Ala
    85          90          95

Arg Ser Ala Arg Leu Ala Cys Arg Arg Leu Pro Ala Ser Val Pro Thr
    100          105          110

```

204

Thr Arg Arg Asp Pro Arg Glu Arg  
115 120

## (2) INFORMATION FOR SEQ ID NO:191:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 89 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:191:

Leu Ala Cys Glu Cys His Arg Arg Tyr Asp Val Gly Ile Glu Phe Arg  
1 5 10 15

Gly Pro Ala Gly Pro Val Ala Thr Gln Ser Gly Pro Pro Gly Pro Ser  
20 25 30

Ile Ala Glu Gly Arg Glu Val Arg Ala Gln Cys Gly Ala Gly Phe Leu  
35 40 45

Glu Arg Arg Pro Ala Val Ser Gly Ala Leu Pro Pro Asn Asn Ala Ser  
50 55 60

Pro Gly Ile Arg Ser Arg Ala Ala Asp Ser Gln Arg His Leu Leu Ala  
65 70 75 80

Gly Asp Gly Ser Asp Val Thr Val Gly  
85

## (2) INFORMATION FOR SEQ ID NO:192:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 119 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:192:

Ala Ser Leu Leu Ala Tyr Ser Ala Ala Ala Ser Thr Ala Leu Ala  
1 5 10 15

Val Ala Cys Val Arg Ala Asp His Arg Asp Arg Arg Thr Ile Arg Asp  
20 25 30

205

```

His Leu Ala Met Ile His Leu Ala Gln Leu Val Thr Gln Pro Pro Gly
  35                               40                               45

Gly Val Arg Gln Arg Leu His His Leu Gly Ile Ala Val Ala Pro Gln
  50                               55                               60

Pro Gln Gln Val Val Val Leu Ala His His Leu Val Thr Gly Thr Gly
  65                               70                               75                               80

Glu Val Gln Gly Glu Gly Arg His Val Ala Ala Glu Val Val Asp Pro
  85                               90                               95

Glu Asn Gln Ile Leu Arg Gln Val Leu Gly Pro Ala Pro His Asp Lys
 100                               105                               110

Pro Asp Ala Gly Ile Gly Gln
 115

```

## (3) INFORMATION FOR SEQ ID NO:193:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 116 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:193:

```

Arg Ala Arg Gly His Arg Ser Ser Lys Gly Ser Arg Trp Ser His Glu
1           5           10           15

Val Leu Glu Gly Cys Ile Leu Ala Asp Ser Arg Gln Ser Lys Thr Ala
  20           25           30

Ala Ser Pro Ser Pro Ser Arg Pro Gln Ser Ser Ser Asn Asn Ser Val
  35           40           45

Pro Gly Ala Pro Asn Arg Val Ser Phe Ala Lys Leu Arg Glu Pro Leu
  50           55           60

Glu Val Pro Gly Leu Leu Asp Val Gln Thr Asp Ser Phe Glu Trp Leu
  65           70           75           80

Ile Gly Ser Pro Arg Trp Arg Glu Ser Ala Ala Glu Arg Gly Asp Val
  85           90           95

Asn Pro Val Gly Gly Leu Glu Glu Val Leu Tyr Glu Leu Ser Pro Ile
 100           105           110

Glu Asp Phe Ser
 115

```

## (2) INFORMATION FOR SEQ ID NO:194:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 911 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:194:

```

TGCTACGCAG CAATCGCTTT GGTGACAGAT GTGGATGCCG GCTGCGCTGC TGGCGATGGC      60
GTGAAAGCCG CCGACGTGTT CCGCGCATTC GGGGAGAACA TCGAACTGCT CAAAAGGCTG      120
GTGCGGGCCG CCGTCGATCG GGTCCCGGAC GAGCGCACGT GCACGCACTG TCAACACCCG      180
GCCGCTTCTC CGTTGCGGTT CGAGCTGCCA TGAGGGTGCT GCTGACCGGC GCGGCGGGCT      240
TCATCGGGTC GCGCGTGCGT GCGGCGTTAC GGGCTGCGGG TCACGACGTG GTGGGCGTCC      300
ACGCGCTGCT GCGCGCGCGC CACCGGCCAA ACCCGTCTCT GCGACCGGAC TGCGAGCGGG      360
TCGACGTGCG CGACCGCAGC GCGCTGGGCG CGTGTGTTGC CGTGTGCGAT CTGCTGTGTC      420
ACCAAGCCGC CATGCTGGGT GCGGGGCTCA ACCGCGCGCA CGCAGCGGCC TATGGCGGCG      480
ACAAGGATTT CCGCACTAGC GTGCTGCTCG CGCGATCTTT CGCGCGCGGG GTCCGCGGTT      540
TGGTGGTGCG GTGCTCGATG GTGCTTTAGG GCGAGGGGCG CTATGACTGT CCGCAGCATG      600
GACCGGTGCA CCGCTGCGCG CGCGCGCGAG CGGACCTGGA CAATGCGGTC TTCGAGCACC      660
GTGCGCGCGG GTGCGCGCGG CCGTCATCTT GCGAATTGGT CGACGAAGAT GCGCGGTTGC      720
GCGCGCGCGG CCGGTACCGG GCAGCAAGAC CGCGCAGGAG CACTACCGGC TGGCTGTGTC      780
GGAAACGAAT GCGGCTTCCG TGGTGGGCTT G

```

## (2) INFORMATION FOR SEQ ID NO:195:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 966 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:195:

GTCCCGCGAT	GTGGCCGAGC	ATGACTTTCC	GCAACACCGG	CGTAGTAGTC	GAAGATATCG	60
GACTTTGTGG	TCCCGGTGGC	GGGATAGAGC	ACCTGTGGGC	GTTGGTCAGC	GTCACCCCCT	120
GCTCGGACGC	CGAACCCATG	CTTTCACCGT	AGCCTGTCCG	TCACACCACT	CSCGAGCGTA	180
ACGTCACGGT	CAAAATATGC	GTGGAAATTC	GCCGTGACCT	TCCGCTCGCG	GACAAATCAAG	240
GCATACTCAC	TTACATCCGA	GCCATTTCGA	CGGCTTCGAT	CGCCTTCGGG	CTGGTCGAAC	300
TGCCGGTCAA	GCTGTACAGC	GCTACCCGAG	ACCACGACAT	CAGCTTCAC	CAGGTGCACG	360
CCAGGACAA	CGGACGCATC	CGGTACAGC	GGCTCTGCGA	GGCGTGTGGC	GAGGTGCTCC	420
ACTACCGCGA	TCTTCCCGG	GCTTACGAGT	CGGGCGAGCG	CCAAATGGTG	GCGATCACCG	480
ACGACGACAT	CCTCAGCTTG	CCTGAAGAAC	GCACCCGGGA	GATCGAGGTG	TTGGAGTTCC	540
TCCCGCGCGC	CGAGCTGGAC	CGGATGATCT	TGACCCGACG	CTACTTTTTG	GAGCCTGATT	600
CGAAGTCCTC	GAAATCCTAT	GTGCTGCTGG	CTAGACACT	CGCCGAGACC	GACCGGATGG	660
CGATCGTCCA	TGCCCCCACC	GGCCGTGAAT	GCAGGAAAAA	TAGAGCCCGC	TATCCACAAAT	720
TGGCGTCCA	GCTCGGCTAC	CACAAACCGT	AGAACHATCG	AGACATTCCG	GAGCTGAAGT	780
GCGGCGCTAT	AGAAGCCGCT	CTGCGCGATT	ATCAACGCGA	AAATACGCTT	ACTCATGCCA	840
TGGCGCTGC	TCACCCCGATG	CGACSTTTTT	GCCACGCTCC	ACCCCTGCGC	GCGCGACCTC	900
AACTGGGCGT	GCATCCGACC	CGTCCCGCGA	AACCGGTTCG	GGCGGGTCCG	CTCATCCCTT	960
CATCCT						966

## (12) INFORMATION FOR SEQ ID NO:196:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2387 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:196:

CCGCACCCGC	GGCAATACCG	CCAGCGGCAC	CGTACCGGCC	GTTTGGGCGG	TTGCCCCCGT	60
TGCGGCCCCG	CCCCCGGGC	CGCCGATGG	AGTTCTCATC	CCCAAAAGTA	CTGGCGTTGC	120
CACCGGAGCC	GGCGTTGCCG	CCGTACCGGC	CAGCCCCGCG	GACTCCACCG	GGCCCCCGGA	180



CTCCGCCGCT	GCCACCGTTC	CCGCCGTTGC	CGATCAACAT	GCCGCTGGCG	CCACCCCTTGC	240
CACCCACGCC	ACCGGCTCCG	CCCACCCGSC	CGACACCAAG	CGAGCTGCGG	CCGGAGCCAC	300
CATCACCACC	TACGCCACCG	ACCGCCACGA	CACCAAGGAC	CGGCTTTTCG	TGAAACGTGG	360
CGGTGCCACC	ACCGCCGCGG	TTACCGCCAA	CCCCACCGGC	AACGCCGGCG	CCGCCATCCC	420
CGCCGCCCGC	GGCGTTGCGG	CCGTTGCCGC	CGTTGCCGAA	CAACAACCGG	CCGGCGCCGC	480
CGTTGCCGCG	CGCGCCGCGG	GTCCCGCGCG	CGCGCGCGAC	GCCAGGCGCG	CTGCCGCGCT	540
TGCGGCCATC	ACCACCGTTC	CCGCCGACCA	CATCGGGTTC	TGCTTCGGGG	TCTGGGCTGT	600
CAACCTCGC	GATGCCAGCG	TTGCCGCGCG	TTCCCGCGGG	CCCCCGCGTG	GCGCGGTAC	660
CACGATACC	ACCGCGCCCA	CGCGCGCCAC	CGTTGCCGCG	ATCACCGAAT	AGCAACCGGC	720
CGCGCGCACC	ATTGCCGCGA	GCTCCCGCTG	CGCGCGCTC	GGCGCGCGGG	GCGGCACTGG	780
CAGCCCCGTT	ACCACCGAAA	CGCGCGCTAC	CACCGGTAGA	GGTGGCAGTG	CGCATGTGTA	840
CGAAAGCGCC	GCTTCGCGCG	CGCGCGCTAC	CACCGCGACT	GCGCGCGCGT	ACACCGTGGG	900
ACCGCTTGGC	ACCATCACCG	CGAAAGCGGC	TGCAATGTC	GCGCTGCGCG	ACTCGCGCGT	960
CGCCCGCGTT	GCGCGCGCGG	CCACCGCGAG	CGCGCGGTACC	GCGGTACCA	CGCGCAGCGC	1020
CGGTGCGCTT	GCGCGAGCGT	GCGGTGCGCG	TGCGACCGTC	GCGCGCGCGT	CCACCGGTGG	1080
GCGTGCAGCG	AGTGCATGG	CGCGCGGTGC	CGCGGTGCGC	GCGGTTTGA	TCACCGATGC	1140
CGGACACATC	TGCGCGGCTG	TCCCGGTGC	TGCGCGCGGG	GCGCGGCGTG	GGATTGACCC	1200
CGTTTGCGCC	GGCGAGGCGG	GCGCGCGCGG	TACACCGGCG	GCGCGCATGG	CCGAACAGCC	1260
CGCGCTTGGC	GCGGTACCG	CGCGCACCGC	CGATGCGTGC	GCGCACGCTG	GTCCCGCGCA	1320
CACCGCGCTT	GCGCGCGGTG	CCCCACAACT	ACCGCGCGCTT	CCACCGCGCA	CGCGCGCGCG	1380
CGCGCGTACC	ACCGCGCGCG	CGGTGCGCGC	CGTTGCGCAT	CACCGCGCGC	GCGCGTCCGC	1440
TGCGCGCGGT	TTGACCGAAC	CGCGAGCGCG	CGCGGTTCGC	ACCGTTGCGA	AACAGCAACC	1500
CGCGCGCGCG	GCGAGGCTGC	CGGGGTGCGG	TCCCGTGGCG	GCGGTTCGCG	ATCAACGGGC	1560
GCGCGCGCGG	CGCGTGGTG	GGCGCATTCG	CGCGACCGAG	CAGACTGCGC	TCAACAGCGG	1620
CTTCAGTGCT	GCGATACCGA	CGCGCGCGCG	CAGTCAACGC	CTCAGCAAA	TGCTCGTGAA	1680
AGCGTGCAC	CTGTACGCTG	AGCGCGTGTG	ACTGCGAGCG	ATGGGCGCGG	AACAACCGCG	1740
CAATCGCGCG	CGACACTTCA	TGCGCAGCGG	CAGCGACCGC	TTCCGTGCTC	GGGATCGCGG	1800

```

CGGCCGCATT AGCCGCGCTC ACCTGCSAAC CAATAGTCGA TAAATCGAAA GCGCGCASTTG      1860
CCAGCAGCTG CCGCTGCGCG ATCACCAGAG ACACCTGCGA CCTCCGATA CCCCATATCG      1920
CGGCACCTG TCCCCAGCG CCAGGTGACC TTTGGTGGCT GCGTGGGGGC CCTGACTATG      1980
CCCCCGACGG CCTCTGTTCT GATTGCCCCC GCGCGCGAGC TTGTTGCGCG AGTTGAGAC      2040
GGGAGGACAG GCGAGCTTG GTGTAGACCT GGGTCAAGTG GGAATGACG GTCCGCGCGC      2100
AGATGAATAG GCGGACGCG ATCTCCTTCT TGTGAGTCC CTCACCGACC AGTAGAGCCA      2160
CCTCAAGCTC TGTGCGTGT AACGCGCCCC AGCCACTTCT CCGGCGTTTC CGTGACCGCG      2220
GGCTCGTTT CCGGTACGCG ATCGGCTCAT CGATCGATAA CCGAGTTCTT TCGGCCGAGG      2280
CATCGTCGAA CTCGCTGTCA CCCATGGATT TCGAAGGCT CGCTAGCGAC GAGTTACAGC      2340
CGGCTGGTA GATCCCGAAG CGGACCG                                     2367

```

## (2) INFORMATION FOR SEQ ID NO:197:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 376 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:197:

```

Gln Pro Ala Gly Ala Thr Ile Ala Ala Ser Ser Pro Cys Ala Thr Val
1           5           10           15

Gly Ala Gly Gly Gly Thr Gly Ser Pro Val Thr Thr Glu Thr Ala Ala
20          25          30

Thr Thr Gly Arg Gly Gly Ser Gly Asp Val Tyr Glu Ser Ala Ala Ser
35          40          45

Gly Ala Ala Ala Thr Thr Pro Thr Ala Gly Gly Tyr Thr Val Gly Pro
50          55          60

Val Ala Thr Ile Thr Ala Lys Gly Ala Arg Asn Val Ala Leu Arg Asp
65          70          75          80

Ser Ala Val Ala Ala Val Ala Ala Ala Ala Thr Gly Ser Gly Gly Thr
85          90          95

Ala Val Thr Thr Gly Thr Ala Gly Gly Leu Ala Arg Ala Cys Arg Arg
100         105         110

```

210

Gly Gly Thr Val Ala Ala Gly Ala Thr Gly Arg Arg Ala Gly Ser Ala		
115	120	125
Met Ala Ala Arg Ala Ala Val Ala Ala Gly Leu Ile Thr Asp Ala Gly		
130	135	140
His Ile Cys Arg Ala Val Pro Gly Ala Gly Arg Gly Ala Gly Arg Gly		
145	150	155
Ile Asp Pro Val Cys Pro Gly Glu Ala Gly Ala Ala Gly Thr Thr Gly		
165	170	175
Ala Ala Met Ala Glu Gln Pro Gly Val Ala Ala Val Thr Ala Arg Thr		
180	185	190
Pro Asp Ala Cys Gly His Ala Gly Ala Ala Asp Thr Ala Val Ala Ala		
195	200	205
Val Ala Pro Gln Pro Pro Pro Val Pro Thr Gly Thr Ala Gly Arg Ala		
210	215	220
Gly Thr Thr Gly Pro Ala Val Ala Ala Val Ala Asp Gln Pro Gly Arg		
225	230	235
Ala Ser Ala Ala Ala Gly Leu Thr Glu Pro Ala Ser Arg Ala Val Ala		
245	250	255
Thr Val Ala Lys Gln Gln Pro Ala Gly Arg Ala Arg Leu Pro Gly Cys		
260	265	270
Arg Pro Val Gly Ala Val Ser Asp Gln Arg Ala Pro Gln Lys Arg Leu		
275	280	285
Gly Gly Arg Ile His Arg Thr Gln Gln Thr Pro Leu Asn Ser Gly Phe		
290	295	300
Ser Ala Gly Ile Pro Thr Arg Gly Arg Ser Gln Arg Leu His Lys Leu		
305	310	315
Leu Val Lys Arg Cys His Leu Tyr Ala Glu Arg Leu Ile Leu Pro Ser		
325	330	335
Met Gly Pro Glu Gln Pro Arg Asn Arg Arg Arg His Phe Ile Gly Ser		
340	345	350
Arg Ser His His Phe Arg Arg Arg Asp Arg Arg Gly Arg Ile Ser Arg		
355	360	365
Ala His Leu Arg Thr Asn Ser Arg		
370	375	

(2) INFORMATION FOR SEQ ID NO:198:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2852 base pairs

(B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:198:

GGCCAAACCG	CCCCGGCGAT	CGCCGCCACC	GAGGCGCCCT	ACGACCAGAT	GTGGGCCCAAG	60
GACGTGGGCG	CGATGTTTGG	CTACCATGCC	GGGGCTTCGG	CGGCGGTCTC	GGCGTTGACA	120
CGTTTGGGCC	AGGCGCTGCC	GACCGTGGCG	GGGCGCGGTC	CGCTGGTCAG	CGCGGGCGCG	180
GCTCAGGTGA	CCACGCGGGT	CTTCGCCAAC	CTGGGCTTGG	CGAACGTCCG	CGAGGGCAAC	240
GTCCGCCAAC	GTAATGTCCG	GAAGTTCAAT	CTGGGCTTCG	CCAACATCGG	CAACGGCAAC	300
ATCGGCAGCG	GCAACATCGG	CAGCTCCAAC	ATGGGTTTTC	GCAACGTGGG	TCTGGGTTTC	360
ACCGCAGCGC	TGACCAACAT	CGGTTTGGGC	AACACCGGCG	GCAACACAT	CGGTTTGGGC	420
AACACCGGCG	GCAACACAT	CGGTTTGGGC	AATACCGGAG	ACGCGAACCG	AGGTATGGCG	480
CTCAGCGGTA	CGGTTTGGT	GGGTTTGGGC	GGGCTGAAT	CGGGCACCGG	CAACATCGGT	540
CTGTTCAACT	CGGGCACCGG	AAACGTGGGC	ATCGGCAACT	CGGGTACCGG	GAAGTTGGGC	600
ATTGSCAACT	CGGGCAACAG	CTACAACACC	GGTTTGGGCA	ACTCGGCGCA	CGCCAACACC	660
GGCTTCTTCA	ACTCGGCAAT	AGCCAACACC	GGGCTGGGCA	ACCGCGGCAA	CTACAACACC	720
GGTAGCTACA	ACCGCGGCAA	CAGCAATACC	GGGCGCTTCA	ACATGGGCGA	GTACAACACC	780
GGCTACCTGA	ACAGCGGCAA	CTACAACACC	GGCTTGGCAA	ACTCGGCGCA	TGTCAACACC	840
GGCGCCTTCA	TACTGGGCAA	CTTCAACAAC	GGCTTCTTGT	GGCGGGGCGA	CCACCAAGGC	900
CTGATTTTCT	GGAGCCCGCG	CTTCTTCAC	TGGACCACTG	CGCCGTGGTC	GGGATTTCTC	960
AACAGCGGTG	CGGTAGGCG	GTCCGGCTTC	CTGAATTCGG	GTCGCAACAA	TTCTGGCTTC	1020
TTCAACTCTT	CGTGGGGGCG	CATCGGTAAAC	TCGGGCTTGG	CAAACGGGGG	CGTCTGGGTA	1080
TGGGGCGTGA	TCAACTGGGG	CAACACCGTA	TGGGTTTGT	TCAACATGAG	CTGGGTGGCC	1140
ATCACACCGC	CGGCTTGAAT	CTCGGGCTTC	TTCAACACCG	GAAGCAACAT	GTGGGGATTT	1200
TTGGGTGGCC	CAGGGGTCTT	CAATCTCGGC	CTGGCAAAAC	GGGGGCTCGT	GAACATTCTC	1260
GGCAACGGCA	ACATCGGCAA	TTACAACATT	CTGGGCGCGG	GAAACGTGGG	TGACTTCAAC	1320
ATCTTGGGCA	CGGGCAACCT	CGGCAGCCAA	AACATCTTGG	GCAGCGGCAA	CGTGGGCGCG	1380

TTCAATATCG	GCAGTGGAAA	CATCGGAGTA	TTCAATGTCTG	GTTCGCGAAG	CCTCGGAAAC	1440
TACAACATCG	GATCGGGAAA	CCTCGGGATC	TACACATCTG	GTTTGGGAAA	CGTCGGCGAC	1500
TACAACGTCTG	GCTTCGGGAA	CSCGGGCGAC	TTCAACCAAG	GCTTTGCCAA	CACCGGCAAC	1560
AACAACATCG	GGTTCGCCAA	CACCGGCAAC	AACAACATCG	GCATCGGGCT	GTCCGGCGAC	1620
AACCAGCAGG	GCTTCAATAT	TGCTAGCGGC	TGGAACTCGG	GCACCGGCAA	CAGCGGGCTG	1680
TTCAATTCTG	GCACCAATAA	GTTTGGCATC	TTCAACGCCTG	GCACCGGAAA	CGTCGGCATC	1740
GCAAACTCTG	GCACCGGAAA	CTGGGGTATC	GGGAACCGCG	GTACCGGAAA	TACCGGCATC	1800
CTCAATGCTG	GCAGCTACAA	CACGGGCTATC	CTCAACGCGG	GGGACTTCAA	CACGGGCTTC	1860
TACAACACCG	GCAGCTACAA	CACGGGCTATC	TTCAACGTCTG	GTACACCAAA	CACCGGCAAC	1920
TTCAACGTCTG	GTACACCAAA	TACCGGCAAC	TATAACCGCG	GTACACCAAA	CACCGGCTTC	1980
TTCAATCCCG	GCACCTTCAA	TACCGGCTATC	TTCAACGTCTG	GGGACTTCAA	CAATGGCTTC	2040
TTGGTGGCGG	GCACCTTCAA	GGGCTAGATC	GGGCTAGATC	TCTGGCTCAC	CATCTCATTC	2100
ATCCCCATAA	ACGAGCAGAT	GGTCAATTGC	GTACACCAAC	TATGAGCTTC	CGGCGGCAAC	2160
ATGATCAGCG	TCACCGAGGC	CTCGACCGTT	TTCCCCCAAA	CCTTCTATCT	GAGCGGTTTG	2220
TCTTCTCTCG	GGCGGCTCAA	TCTCAGCGCA	TCCACGCTTA	CGTTCGGAC	GATCAGCTTC	2280
ACCATCGGCG	GACCGAGCGT	GACCGTCCCG	ATCAGCATTC	TGCTGCTCT	GGGAGCGCGC	2340
ACCATTACCT	TCTCAAGAT	CGATCGGGCG	CGGGGCTATC	GAAATTCGAC	CACCAACCGC	2400
TCTTCCGGCT	TCTTCAACTC	GGGACCGCGT	GGGACATCTG	GCTTCCAAA	CGTGGGCGCG	2460
GGCATTTCTG	GCTTCTGAAA	CAGTGGTTTG	AGCAGCGCGA	TAGGGATTC	GGGTTTCCAG	2520
AACCTCGGCT	CGTTCGAGTC	AGGCTGGGCG	AACCTGGGCA	ACTCGGATTC	GGGCTTTTTC	2580
AACACCACTA	CGGTGAACCT	CTCCAGCGCG	GGCATGTCT	CGGGGCTGAA	CAACATCGGC	2640
ACCAACCTGT	CGGCGTGTTC	CGGCGTGTTC	ACCGGACGAA	TTTTCAGCGC	GGGCTTTGDC	2700
AACCTGGGCG	AGTTGAACAT	CGGAGCGCGC	TCTGCGCGAA	TTGGGACGAA	GTTAGATAGC	2760
GTTTCAACAA	TCAATCCCG	GTTTGGCGCG	AGTGCATCAG	ACGAATCGAA	CCCGGGAAGC	2820
GTAACCGAAT	AAACCGAATG	GGGCGCTCTC	AT			2853

(2) INFORMATION FOR SEQ 10 NO:199:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 943 amino acids

213

(B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:199:

Gly	Gln	Asn	Ala	Pro	Ala	Ile	Ala	Ala	Thr	Gln	Ala	Ala	Tyr	Asp	Gln	1	5	10	15
Met	Trp	Ala	Gln	Asp	Val	Ala	Ala	Met	Phe	Gly	Tyr	His	Ala	Gly	Ala	20	25	30	
Ser	Ala	Ala	Val	Ser	Ala	Leu	Thr	Pro	Phe	Gly	Gln	Ala	Leu	Pro	Thr	35	40	45	
Val	Ala	Gly	Gly	Gly	Ala	Leu	Val	Ser	Ala	Ala	Ala	Ala	Gln	Val	Thr	50	55	60	
Thr	Arg	Val	Phe	Arg	Asn	Leu	Gly	Leu	Ala	Asn	Val	Arg	Gln	Gly	Asn	65	70	75	80
Val	Arg	Asn	Gly	Asn	Val	Arg	Asn	Phe	Asn	Leu	Gly	Ser	Ala	Asn	Ile	85	90	95	
Gly	Asn	Gly	Asn	Ile	Gly	Ser	Gly	Asn	Ile	Gly	Ser	Ser	Asn	Ile	Gly	100	105	110	
Phe	Gly	Asn	Val	Gly	Pro	Gly	Leu	Thr	Ala	Ala	Leu	Asn	Asn	Ile	Gly	115	120	125	
Phe	Gly	Asn	Thr	Gly	Ser	Asn	Asn	Ile	Gly	Phe	Gly	Asn	Thr	Gly	Ser	130	135	140	
Asn	Asn	Ile	Gly	Phe	Gly	Asn	Thr	Gly	Asp	Gly	Asn	Arg	Gly	Ile	Gly	145	150	155	160
Leu	Thr	Gly	Ser	Gly	Leu	Leu	Gly	Phe	Gly	Gly	Leu	Asn	Ser	Gly	Thr	165	170	175	
Gly	Asn	Ile	Gly	Leu	Phe	Asn	Ser	Gly	Thr	Gly	Asn	Val	Gly	Ile	Gly	180	185	190	
Asn	Ser	Gly	Thr	Gly	Asn	Trp	Gly	Ile	Gly	Asn	Ser	Gly	Asn	Ser	Tyr	195	200	205	
Asn	Thr	Gly	Phe	Gly	Asn	Ser	Gly	Asp	Ala	Asn	Thr	Gly	Phe	Phe	Asn	210	215	220	
Ser	Gly	Ile	Ala	Asn	Thr	Gly	Val	Gly	Asn	Ala	Gly	Asn	Tyr	Asn	Thr	225	230	235	240

214

Gly Ser Tyr Asn Pro Gly Asn Ser Asn Thr Gly Gly Phe Asn Met Gly	245	250	255
Gln Tyr Asn Thr Gly Tyr Leu Asn Ser Gly Asn Tyr Asn Thr Gly Leu	260	265	270
Ala Asn Ser Gly Asn Val Asn Thr Gly Ala Phe Ile Thr Gly Asn Phe	275	280	285
Asn Asn Gly Phe Leu Trp Arg Gly Asp His Gln Gly Leu Ile Phe Gly	290	295	300
Ser Pro Gly Phe Phe Asn Ser Thr Ser Ala Pro Ser Ser Gly Phe Phe	305	310	315
Asn Ser Gly Ala Gly Ser Ala Ser Gly Phe Leu Asn Ser Gly Ala Asn	325	330	335
Asn Ser Gly Phe Phe Asn Ser Ser Ser Gly Ala Ile Gly Asn Ser Gly	340	345	350
Leu Ala Asn Ala Gly Val Leu Val Ser Gly Val Ile Asn Ser Gly Asn	355	360	365
Thr Val Ser Gly Leu Phe Asn Met Ser Leu Val Ala Ile Thr Thr Pro	370	375	380
Ala Leu Ile Ser Gly Phe Phe Asn Thr Gly Ser Asn Met Ser Gly Phe	385	390	395
Phe Gly Gly Pro Pro Val Phe Asn Leu Gly Leu Ala Asn Arg Gly Val	405	410	415
Val Asn Ile Leu Gly Asn Ala Asn Ile Gly Asn Tyr Asn Ile Leu Gly	420	425	430
Ser Gly Asn Val Gly Asp Phe Asn Ile Leu Gly Ser Gly Asn Leu Gly	435	440	445
Ser Gln Asn Ile Leu Gly Ser Gly Asn Val Gly Ser Phe Asn Ile Gly	450	455	460
Ser Gly Asn Ile Gly Val Phe Asn Val Gly Ser Gly Ser Leu Gly Asn	465	470	475
Tyr Asn Ile Gly Ser Gly Asn Leu Gly Ile Tyr Asn Ile Gly Phe Gly	485	490	495
Asn Val Gly Asp Tyr Asn Val Gly Phe Gly Asn Ala Gly Asp Phe Asn	500	505	510
Gln Gly Phe Ala Asn Thr Gly Asn Asn Asn Ile Gly Phe Ala Asn Thr	515	520	525
Gly Asn Asn Asn Ile Gly Ile Gly Leu Ser Gly Asp Asn Gln Gln Gly			

530	535	540
Phe Asn Ile Ala Ser Gly Trp Asn Ser Gly Thr Gly Asn Ser Gly Leu		
545	550	555
Phe Asn Ser Gly Thr Asn Asn Val Gly Ile Phe Asn Ala Gly Thr Gly		
	565	570
		575
Asn Val Gly Ile Ala Asn Ser Gly Thr Gly Asn Trp Gly Ile Gly Asn		
	580	585
		590
Pro Gly Thr Asp Asn Thr Gly Ile Leu Asn Ala Gly Ser Tyr Asn Thr		
	595	600
		605
Gly Ile Leu Asn Ala Gly Asp Phe Asn Thr Gly Phe Tyr Asn Thr Gly		
	610	615
		620
Ser Tyr Asn Thr Gly Gly Phe Asn Val Gly Asn Thr Asn Thr Gly Asn		
	625	630
		635
Phe Asn Val Gly Asp Thr Asn Thr Gly Ser Tyr Asn Pro Gly Asp Thr		
	645	650
		655
Asn Thr Gly Phe Phe Asn Pro Gly Asn Val Asn Thr Gly Ala Phe Asp		
	660	665
		670
Thr Gly Asp Phe Asn Asn Gly Phe Leu Val Ala Gly Asp Asn Gln Gly		
	675	680
		685
Gln Ile Ala Ile Asp Leu Ser Val Thr Thr Pro Phe Ile Pro Ile Asn		
	690	695
		700
Gln Gln Met Val Ile Asp Val His Asn Val Met Thr Phe Gly Gly Asn		
	705	710
		715
Met Ile Thr Val Thr Glu Ala Ser Thr Val Phe Pro Gln Thr Phe Tyr		
	725	730
		735
Leu Ser Gly Leu Phe Phe Phe Gly Pro Val Asn Leu Ser Ala Ser Thr		
	740	745
		750
Leu Thr Val Pro Thr Ile Thr Leu Thr Ile Gly Gly Pro Thr Val Thr		
	755	760
		765
Val Pro Ile Ser Ile Val Gly Ala Leu Glu Ser Arg Thr Ile Thr Phe		
	770	775
		780
Leu Lys Ile Asp Pro Ala Pro Gly Ile Gly Asn Ser Thr Thr Asn Pro		
	785	790
		795
Ser Ser Gly Phe Phe Asn Ser Gly Thr Gly Gly Thr Ser Gly Phe Gln		
	805	810
		815
Asn Val Gly Gly Gly Ser Ser Gly Val Trp Asn Ser Gly Leu Ser Ser		
	820	825
		830



216

Ala	Ile	Gly	Asn	Ser	Gly	Phe	Gln	Asn	Leu	Gly	Ser	Leu	Gln	Ser	Gly	835	840	845	
Trp	Ala	Asn	Leu	Gly	Asn	Ser	Val	Ser	Gly	Phe	Phe	Asn	Thr	Ser	Thr	850	855	860	
Val	Asn	Leu	Ser	Thr	Pro	Ala	Asn	Val	Ser	Gly	Leu	Asn	Asn	Ile	Gly	865	870	875	880
Thr	Asn	Leu	Ser	Gly	Val	Phe	Arg	Gly	Pro	Thr	Gly	Thr	Ile	Phe	Asn	885	890	895	
Ala	Gly	Leu	Ala	Asn	Leu	Gly	Gln	Leu	Asn	Ile	Gly	Ser	Ala	Ser	Cys	900	905	910	
Arg	Ile	Arg	His	Glu	Leu	Asp	Thr	Val	Ser	Thr	Ile	Ile	Ser	Ala	Phe	915	920	925	
Cys	Gly	Ser	Ala	Ser	Asp	Glu	Ser	Asn	Pro	Gly	Ser	Val	Ser	Glu		930	935	940	

(2) INFORMATION FOR SEQ ID NO:200:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 53 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(\*1) SEQUENCE DESCRIPTION: SEQ ID NO:200:

GGATCCATAT GGGCCATCAT CATCATCATC ACGTGATCGA CATCATCGGG ACC

53

(2) INFORMATION FOR SEQ ID NO:201:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 42 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(\*1) SEQUENCE DESCRIPTION: SEQ ID NO:201:

GCTGAATTCA GGCTCGGTT GCGCCGCCCT CATCTTGAAC GA

42

(2) INFORMATION FOR SEQ ID NO:202:

(1) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 31 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:202:

GGATCCTGCGA GCGTCGAAGC CACCGAGCGG T 31

(2) INFORMATION FOR SEQ ID NO:203:

(1) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 31 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:203:

CTCTGAATTC AGCGCTGGAA ATCGTCGCGA T 31

(2) INFORMATION FOR SEQ ID NO:204:

(1) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 33 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:204:

GGATCCAGCG CTGACATGAA GACCGATGCC GGT 33

(2) INFORMATION FOR SEQ ID NO:205:

(1) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 38 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:205:

GGATATCTGC AGAATTCAGG TTAAAGCCC ATTTCCGA

38

(2) INFORMATION FOR SEQ ID NO:206:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:206:

CCGCATGCCA GCCACCTGCC CACAACGGCC

36

(2) INFORMATION FOR SEQ ID NO:207:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 37 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:207:

CTTCATGGA TTCTCAGGCC GGTAAAGTCC GCTGCGG

37

(2) INFORMATION FOR SEQ ID NO:208:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 7676 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:208:

TGGCGATCG GACGCGCCCT GTAGCGCCGC ATTAAGCCCG GCGGCTCTCG TGCTTACCG

60

CAGCGTGACC	GCTACACTTG	CCAGCGCCCT	AGCGCCCGCT	CGTTTCGCTT	TCTTCCCTTC	120
CTTTCTCGCC	ACGTTTCGCC	GCCTTCGCCG	TCAAGCTCTA	AATCGGGGGC	TCCCTTTAGG	180
GTTCGATTT	AGTGCTTTAC	GGCACCTCGA	CCCCAAAAA	CTTGATTAGG	GTGATGGTTC	240
ACGTAGTGGG	CCATCGCCCT	GATAGAGGCT	TTTTCGCCCT	TTGACGTTGG	AGTCCACCTT	300
CTTTAATAGT	GGACTCTTGT	TCCAAACTGG	AACAACACTC	AACCCATCTT	CGCTCTATTC	360
TTTTCGATTTA	TAAGGGATTT	TGCCGATTTC	GGCCTATTGG	TTAAAAAATG	AGCTGATTTA	420
ACAAAAATTT	AACCGGATTT	TTAACAAAT	ATTACCTTTT	ACAATTTCTG	GTGGCACTTT	480
TGCGGGAAAT	GTGCGCGGAA	CCCCATTTTG	TTTATTTTTT	TAAATACATT	CAAATATGTA	540
TCCGCTCATG	AATTAATTCT	TAGAAAAACT	CATCGAGCAT	CAATGAAAAC	TGCAATTTAT	600
TCATATCAGG	ATTATCATA	CCATATTTTT	GAAAAAGCCG	TTTCTGYAAT	GAAAGAGAAA	660
ACTCAGCGAG	GCAGTTCCAT	AGCATGGCAA	GATCCTGGTA	TGGGTCGCG	ATTCCGACTC	720
GTCCAACATC	AATACAACCT	ATTAAATTC	CCTGCTCATA	AATAAGGTTA	TCAAGTGAGA	780
AATCACCATG	AGTGACGACT	GAATCCCGTG	AGATGGGCAA	AAGTTTATGC	ATTTCTTTCC	840
AGACTTGTTC	AACAGGCCAG	CCATTAGGCT	GCTCATCAAA	ATCACTGGCA	TCAACCAAAC	900
CGTATTCAT	TGCTGATTGC	GCCTGAGCTA	GACGAATAC	GCGATCGCTG	TTAAAAGGAC	960
AATTACAAAC	AGGAATCGAA	TGCACCGGG	GCGGAACAC	TGCCAGGCGA	TCAACAATAT	1020
TTTCAGCTCA	ATCAGGATAT	TCTTCTAATA	CCTGGGAATG	TGTTTTCCCG	GGGATCGCAG	1080
TGCTCACTAA	CCATGCATCA	TCAGGAGTAC	GGATAAAATG	CTTGATGCTC	GGAAGAGGCA	1140
TAAATTCCGT	CAGCCAGTTT	AGTCTGACCA	TCTCATCTGT	AACATCATTG	GCAACGCTAC	1200
CTTTGCCATG	TTTCAGAAAC	AATCTGCGCG	CATCGGGCTT	CCCATACAAT	CGATAGATTG	1260
TGCGACCTCA	TTGCCCCGAC	TTATCGCGAG	CCCATTTATA	CCCATATAAA	TCAGCATCCA	1320
TGTTGCAATT	TAATCGCGGC	CTAGAGCAAG	ACGTTTCCCG	TTGAATATGG	CTCATACAC	1380
CCCTTGATTT	ACTGTTTATG	TAAGCAGACA	GTTTTATTGT	TCATGACCAA	AATCCCTTAA	1440
CGTGAGTTTT	CGTTCCACTG	AGCCTCAGAC	CCCGTAGAAA	AGATCAGAGG	ATCTTCTTGA	1500
GATCCTTTTT	TTCTGCGGCT	AATCTGCTGC	TTGCAACCAA	AAAAACGACC	GCTACCAGCG	1560
GIGGTTTTGT	TGCGGATCA	AGNCTACCA	ACTCTTTTTC	CGAAGGTAA	TGGTTTCAGC	1620
AGAGCGCAG	TACCAATAC	TGTCCTTCTA	GTGTAGCCCT	AGTTAGGCCA	CCACTTCAAG	1680

AACCTCTGTAG	CACCGCCTAC	ATACCTCGCT	CTGCTAATCC	TGTTACCACT	GGCTGCTGCC	1740
AGTGGCGATA	AGTCGTGTCT	TACCGGCTTG	GACTCAAGAC	GATAGTTACC	GCATAAGGCG	1800
CAGCGGTCCG	GCTGAACGGG	GGGTTCTGTC	ACACAGGCCA	GCTTGGAGCG	AACGACCTAC	1860
ACCGAACTGA	GATACCTACA	GGGTGAGCTA	TGAGAAAGCG	CCACGCTTCC	CGAAGGGAGA	1920
AAGGCGGACA	GGTATCCGGT	AAGCGGCAGG	GTCGGAACAG	GAGAGCGCAC	GAGGGAGCTT	1980
CCAGCGGGAA	ACGCCCTGTA	TCTTTATAGT	CCTGTCCGGT	TTCCGCCACT	CTGACTTCAG	2040
CGTCGATTTT	TGTGATGCTC	GTCAGGGGGG	CGAGGCTAT	GGAAAAACGC	CAGCAACGGG	2100
GCCTTTTTC	GGTTCTGCG	CTTTTCTGG	CCTTTTCTC	ACATGTCTCT	TCCTGCGTTA	2160
TCCCTGAT	CTGTGGATAA	CCGTATTACC	GCCTTTGAGT	GAGCTGATAC	CGCTCGCCGC	2220
AGCCGAACCA	CCGAGGCGAG	CGAGTCACTG	AGCGAGGAAG	CGGAAGAGCG	CCTGATGCGG	2280
TATTTTCTCC	TACGCGATCT	GTGCGGTATF	TCACACCGCA	TATATGCTGC	ACTCTCAGTA	2340
CAATCTGCTC	TGATGCCGCA	TAGTTAAGCC	AGTATACACT	CGCTATGCC	TACGTGACTG	2400
GGTCATGGCT	GCGCCCTGAC	ACCGCCCAAC	ACCGGCTGAC	GCGCCCTGAC	GGGCTTGCTC	2460
GCTCCCGGCA	TCCGCTTACA	GACAAGCTGT	GACCGTCTCC	GGGAGCTGCA	TGTGTCAAGG	2520
GTTTTCACCG	TCATCAGCGA	AACGCGCGAG	GCACCTGCGG	TAAAGCTCAT	CAGCCTGGTC	2580
GTGAAGCGAT	TCACAGATGT	CTGCGTGTTC	ATCCGCTCTC	AGCTCCTTGA	GTTTCTCCAG	2640
AAGCCTTAAT	GTCTGGCTTC	TGATAAAGCG	GGCATGTTA	AGGCGGCTTT	TTTCTGTGTT	2700
GGTCACTGAT	GCCTCCGTGT	AAGCGGGATT	TCTGTTGATG	GGGTAAATGA	TACCGATGAA	2760
ACGAGAGAGG	ATGCTCAGCA	TACGGTTTAC	TGATGATGAA	CATGCCCGGT	TACTGGAACG	2820
TTGTGAGGCT	AAACAACCTG	CGTATGGAT	GCGCGGGGAC	CAGAGAAAAA	TCCTCAGGG	2880
TCAATGCCAG	CGCTTCGTTA	ATACAGATGT	AGGTGTTCCA	CAGGCTAGCC	AGCAGCATCC	2940
TGCGATGCAG	ATCGGGAACA	TAATGGTGCA	GGGCGCTGAC	TTCCGCGTTF	CCAGACTTTA	3000
CGAAACACGG	AAACCGAAGA	CCATTGATGT	TGTTGCTCAG	GTGCGAGACG	TTTTGCGGCA	3060
GCAGTCCCTT	CAGTTTCCTT	CGGATATCGG	TGATTGATTC	TGCTAACCAG	TAAGGCAACC	3120
CGGCGAGCTT	AGCGGGGTCC	TCAACGACAG	GAGCAGGATC	ATGCGGACCC	GTGGGCGCGC	3180
CATGCCGGCG	ATAATGGGCT	GCTTCTCGCC	GAAACGTTTT	GTGGCGGAC	CAGTGACGAA	3240
GGCTTGAGCG	AGGGCGTGCA	AGATTCCGAA	TACCGCAAGC	GACAGGCCGA	TCATCGTCGC	3300
GCTCCAGCGA	AAGCGGTCTT	CGCCGAAAT	GACCCAGAGC	GCTGCCGCGA	CCTGTCTTAC	3360

GAGTTGCATG	ATAAAGAAGA	CAGTCATAG	TGGGCGACG	ATAGTCATGC	CCCGCGCCCA	3420
CCGGAAGGAG	CTGACTGGGT	TGAAGGCTCT	CAAGGGCATC	GGTCGAGATC	CCGGTGCCTA	3480
ATCASTGAGC	TAACCTACAT	TAATTCGGTT	GCGCTCACTG	CCCGCTTTCC	AGTCGGGAAA	3540
CCTGTCGTGC	CAGCTGCATT	AATGAATCGG	CCAACTCCCG	GGGAGAGCCG	GTTTGCGTAT	3600
TGGGCGCCAG	GGTGGTTTTT	CTTTTCACCA	GTGAGACGGG	CAACAGCTGA	TTGCCCTTCA	3660
CCGCTTGCCG	CTGAGAGAGT	TGCAGCAAGC	GGTCCACGGT	GGTTTGCCCC	AGCAGGCGAA	3720
AATCCTGTTT	GATGGTGGTT	AACGGCGGGA	TATAACATGA	GCTGTCTTGG	GTATCGTCTT	3780
ATCCCACTAC	CGAGATATCC	GCACCAACGC	GCAGCCCGGA	CTCGGTAATG	GCGCGCATIG	3840
CGCCAGCCGC	CATCTGATCG	TTGGCAACCA	GCATCGCAGT	GGGAACCATG	CCCTCATTCA	3900
GCATTTGCAT	GGTTTGTTGA	AAACCGGACA	TGGCACTCCA	GTGCGCTTCC	CGTTCCGCTA	3960
TGGGCTGAAT	TTGATTGCCA	GTGAGATATT	TATGCCAGCC	AGCCAGACGC	AGACGCGGCG	4020
AGACAGAACT	TAATGGGCCC	GCTAACAGCG	CGATTGCTCG	GTGACCCAAI	CGGACCGAGT	4080
GCTCCAGGCC	CAGTCGGGTA	CGGCTTTCAT	GGGAGAAAAT	AATACTGTTG	ATGGGTGTCT	4140
GCTCAGAGAC	ATCAAGAAAT	AACGCGGGAA	CATTAGTGCA	GGCAGCTTCC	ACAGCAATGG	4200
CATCCTGGTC	ATCCAGCGGA	TAGTTAATGA	TCAGCCCACT	GACGCGTTGC	GGGAGAAGAT	4260
TGTCACCCGC	CGCTTTACAG	GCTTCGACGC	CGCTTCGTTG	TACCATCGAC	ACCACCAAGC	4320
TGGCAGCCAG	TTGATCGGCG	CGAGATTTA	TCGCGCGGAC	AATTTGGGAC	GGCGCGTGCA	4380
GGCCAGAACT	CGAGGTGGCA	ACGCCAATCA	GCAACGACTG	TTTGCCCCCG	AGTTGTGTGT	4440
CCACCGCGTT	GGGAATGTAA	TTGAGCTCGG	CCATCGCCCG	TTCCACTTTT	TCCCGCGTTT	4500
TGCCAGAAAC	GTGGCTGGCC	TGGTTCCACCA	CCGCGGAAAC	GCTCTGTATA	GAGACACCGG	4560
CATACTCTGC	GACATCGTAT	AACCTTACTG	CTTTCACATT	CACCAGCCTG	AATTGACTCT	4620
CTTCCGGGCG	CTATCATGCC	ATACCGCGAA	AGGTTTTGCG	CCATTCGATG	GTGTCCGGCA	4680
TCTGACGCT	CTGCTTATG	CGACTCCTGC	ATTAGGAAGC	AGCCCACTAG	TAGGTTGAGG	4740
CGTTGAGCA	CCGCGCGCGC	AAGGAATGGT	GCATGCAAGG	AGATGGCGCC	CAACAGTCCC	4800
CCGCGCACCG	CGCTTCCGAC	CATACCCAGG	CCGAAACAAG	CGTCATGAG	CCCGAAGTGG	4860
CGAGCCCGAT	CTTCCCATC	GGTGAATGCG	GGGATATAGG	CGCCAGCAAC	CGCACCTGTG	4920
CGCCCGGTGA	TGCGGCGCAC	GATGCGTCCG	GCTTAGAGGA	TCGAGATCTC	GATCCCGCGA	4980

AATTAATAAG	ACTCACTATA	GGGGAATTGT	GAGCGGATAA	CAATTCDDCT	CTAGAAATAA	5040
TTTTGTTTAA	CTTTAAGAAG	GAGATATACA	TATGGGCCAT	CATCATCATC	ATCACGTGAT	5100
CGACATCATC	GGGACCAGCC	CCACATCCTG	GGAACAGGCG	GCSSCGSAGG	CGGTCCAGCG	5160
GGCGCGGGAT	AGGCTCGATG	ACATCCGCGT	CGCTCGGCTC	ATTGAGCGAG	ACATGCGCGT	5220
GGACAGCGCC	GGCAGATCA	CCTACCGCAT	CAAGCTCGAA	GTGTGCTTCA	AGATGAGGCC	5280
GGGCAACCG	AGGGGCTCGA	AACCACCGAG	CGGTTCGCGT	GAAACGGGCG	CCGGCGCGCG	5340
TACTGTGCGG	ACTACCCCGG	CGTCGTGCGC	GGTGAGCTTG	GCGGAGACCG	GTACCACTCT	5400
GGTCTACCGG	CTGTTCAACC	TGTGCGGTCC	GGCTTTTCAC	GAGAGGTATC	CGAACCTCAC	5460
GATCACCGCT	CAGGGCACCG	GTCTGTGTGC	CGGATCGCG	CAGGCGCGCG	CCGGGACGGT	5520
CAACATTGGG	GGCTCGAGCG	CCTATCTCTC	GGAGGTGAT	ATGGCGCGCG	ACAAGGGGCT	5580
GATGAACATC	GGCTAGGCA	TCTCGGTCA	GGAGTCAAC	TACAACTTGC	CCGGAGTGAG	5640
CGAGCACTTC	AAGCTGAACG	GAAAGTCTT	GGCGGCCATG	TACAGGGCA	CCATCAAAAC	5700
CTGGGACGAC	CCGACATCG	CTGGGCTCAA	CCCCGGGCTG	AACCTGCCCC	GCACCGCGGT	5760
AGTTCGCGTG	CACGGCTCGG	ACGGGTCCGG	TGCACCTTTC	TTGTTCACCC	AGTACCTGTC	5820
CAAGCAAGAT	CCGAGGGCT	GGGCAAGTTC	GGCGGGCTTC	GGCAGGACCG	TGACTTCCC	5880
GGCGGTGCGG	GGTGGCTGG	GTGAGAACGG	CAACGGCGGG	ATGGTGACCG	GTGCGGCGGA	5940
GACAGCGGGC	TGGCTGGCT	ATATCGGCAT	CAGCTTCTTC	GACCAGGCCA	GTCAACGGGG	6000
ACTCGGCGAG	CCCCAAGTAG	GCAATAGCTC	TGGCAATTTC	TTGTTGDDCG	ACGCGCAAGG	6060
CATTCAGGCT	GGGGGGGCTG	GCTTCGCATC	GAAACCCCG	GCGAACGAGG	CGATTTGAT	6120
GATCGACGGG	CCCGGCGCGG	ACGGCTACCC	GATCATCAAC	TACGAGTAGG	CCATCGTCAA	6180
CAACCGGCA	AAGGACGGCG	CCACCGCGCA	GACCTTGACG	GCATTTCTGC	ACTGGGCGAT	6240
CACCGACGGC	AACAAGGGCT	GGTTCCTCGA	CCAGGTTTAT	TTCCAGGCGC	TGCCGCGCGC	6300
GGTGTGAAG	TTGTCTGAGG	CCTTGATGCG	GAGGATTTCC	AGGCTGAGAA	TGAGAGCGGA	6360
TGCGGCTACC	CTCGCGGAGG	AGGCGGTAA	TTTCAGGCGG	ATCTCGGCGG	ACCTGAAAAC	6420
CCAGATCGAC	CAGGTGGAGT	CGACGGCAGG	TTGCTTGACG	GGCGAGTGGC	GCGGCGCGGC	6480
GGGACCGGCC	GGCAGGGCGG	GGTGGTGGG	CTTCCAGGAA	GCAGGCAATA	AGCAGAGGCA	6540
GGAACTCGAC	GAGATCTCGA	CGAATATTGG	TGAGGCGGGC	GTCCAACTACT	CGAGGGCGCA	6600
CGAGGAGCAG	CAGCAGGGCG	TGTCTGCGCA	ATGGGCTTTT	GTGCGGACAA	CGGCGGCTTC	6660

```

GCGGCGGTGG ACCGCTGCAG CGCCACCCGC ACCGGGAGCA CTTGTTGCCC CCCCACCAAC 6720
GGCCGGTGGC AACAGGCGGA ATGCCAGGCC GGGGATCCG AACGCAGCAC CTCCGCGGSC 6780
CGACCCGAAC GCACCGCCGC CACCTGTGAT TGCCCCAAGC GCACCCCAAC CTGTCCGAT 6840
CGACAACCCG GTTGGAGGAT TCAGCTTCGC GTGCTTCTT GGTGGGTGG AGTCTGAGCC 6900
CGCCCACTTC GACTACGGTT CAGCACTCCT CAGCAGAAC ACCGGGAGC CGCCATTTC 6960
CGGACAGCCG CGCGCGGTGG CCAATGACAC CCGTATCCTG CTCGGCCGCG TAGACCAAAA 7020
GCTTTACGCC AGCGCGAAG CCACCGACTC CAAGGCTGCG GCGCGGTGG GCTCGGACAT 7080
GGGTGAGTTC TATATGCCCT ACCCGGSCAC CCGGATCAAC CAGGAAACCG TCTCGCTTGA 7140
CGCCAACGGG GTGTCTGGAA GCGCGTCGTA TTACGAAGTC AAGTTCAGCG ATCCGAGTAA 7200
GCCAAGCGC CAGATCTGGA CGGGCGTAAT CGGCTCGCCG GCGGCGAAG CACCGGAGCC 7260
CGGGCCCCCT CAGCGCTGGT TTGTGTAATG GCTCGGAGC GCCACCAAC CGGTGGACAA 7320
GGGCGGGGCC AAGGCGCTGG CCGAATCGAT CCGGCTTTG GTGCCCCCG CGCGGGGCC 7380
GGCACCGGCT CCTGCAGAGC CCGCTCCGCG GCGGCGGCG GCGGCGAAG TCGCTCTTAC 7440
CCGACAGACA CGACACCGC ACCGGACCTT ACCGGCTTGA GAATCTGCA GATATCATC 7500
ACACTGGCGG CGCTCGAGC ACCACCACCA CCACCACTGA GATCCGGCTG CTAACAAAGC 7560
CGGAAAGGAA GCTGAGTTGG CTGCTGCCAC CGCTGAGCAA TAACTAGCAT AAGCCCTTGG 7620
GGCTCTAAA CGGCTCTTGA GGGGTTTTTT GCTGAAAGGA GGAATATAT CCGGAT 7676

```

## (2) INFORMATION FOR SEQ ID NO:209:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 802 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:209:

```

Met Gly His His His His His Val Ile Asp Ile Ile Gly Thr Ser
1           5           10           15

```

```

Pro Thr Ser Trp Glu Gln Ala Ala Ala Gln Ala Val Gln Arg Ala Arg
20           25           30

```



Asp Ser Val Asp Asp Ile Arg Val Ala Arg Val Ile Glu Glu Asp Met  
 35 40 45  
 Ala Val Asp Ser Ala Gly Lys Ile Thr Tyr Arg Ile Lys Leu Glu Val  
 50 55 60  
 Ser Phe Lys Met Arg Pro Ala Glu Pro Arg Gly Ser Lys Pro Pro Ser  
 65 70 75 80  
 Gly Ser Pro Glu Thr Gly Ala Gly Ala Gly Thr Val Ala Thr Thr Pro  
 85 90 95  
 Ala Ser Ser Pro Val Thr Leu Ala Glu Thr Gly Ser Thr Leu Leu Tyr  
 100 105 110  
 Pro Leu Phe Asn Leu Trp Gly Pro Ala Phe His Glu Arg Tyr Pro Asn  
 115 120 125  
 Val Thr Ile Thr Ala Glu Gly Thr Gly Ser Gly Ala Gly Ile Ala Glu  
 130 135 140  
 Ala Ala Ala Gly Thr Val Asn Ile Gly Ala Ser Asp Ala Tyr Leu Ser  
 145 150 155 160  
 Glu Gly Asp Met Ala Ala His Lys Gly Leu Met Asn Ile Ala Leu Ala  
 165 170 175  
 Ile Ser Ala Glu Glu Val Asn Tyr Asn Leu Pro Gly Val Ser Glu His  
 180 185 190  
 Leu Lys Leu Asn Gly Lys Val Leu Ala Ala Met Tyr Glu Gly Thr Ile  
 195 200 205  
 Lys Thr Trp Asp Asp Pro Glu Ile Ala Ala Leu Asn Pro Gly Val Asn  
 210 215 220  
 Leu Pro Gly Thr Ala Val Val Pro Leu His Arg Ser Asp Gly Ser Gly  
 225 230 235 240  
 Asp Thr Phe Leu Phe Thr Glu Tyr Leu Ser Lys Glu Asp Pro Glu Gly  
 245 250 255  
 Trp Gly Lys Ser Pro Gly Phe Gly Thr Thr Val Asp Phe Pro Ala Val  
 260 265 270  
 Pro Gly Ala Leu Gly Glu Asn Gly Asn Gly Gly Met Val Thr Gly Cys  
 275 280 285  
 Ala Glu Thr Pro Gly Cys Val Ala Tyr Ile Gly Ile Ser Phe Leu Asp  
 290 295 300  
 Glu Ala Ser Glu Arg Gly Leu Gly Glu Ala Glu Leu Gly Asn Ser Ser  
 305 310 315 320  
 Gly Asn Phe Leu Leu Pro Asp Ala Glu Ser Ile Glu Ala Ala Ala Ala

225

				325					330				335		
Gly	Phe	Ala	Ser	Lys	Thr	Pro	Ala	Asn	Gln	Ala	Ile	Ser	Met	Ile	Asp
			340					345					350		
Gly	Pro	Ala	Pro	Asp	Gly	Tyr	Pro	Ile	Ile	Asn	Tyr	Glu	Tyr	Ala	Ile
		355					360					365			
Val	Asn	Asn	Arg	Gln	Lys	Asp	Ala	Ala	Thr	Ala	Gln	Thr	Leu	Gln	Ala
		370				375					380				
Phe	Leu	His	Trp	Ala	Ile	Thr	Asp	Gly	Asn	Lys	Ala	Ser	Phe	Leu	Asp
385					390					395					400
Gln	Val	His	Phe	Gln	Pro	Leu	Pro	Pro	Ala	Val	Val	Lys	Leu	Ser	Asp
			405						410					415	
Ala	Leu	Ile	Ala	Thr	Ile	Ser	Ser	Ala	Gln	Met	Lys	Thr	Asp	Ala	Ala
			420					425					430		
Thr	Leu	Ala	Gln	Glu	Ala	Gly	Asn	Phe	Glu	Arg	Ile	Ser	Gly	Asp	Leu
		435					440					445			
Lys	Thr	Gln	Ile	Asp	Gln	Val	Glu	Ser	Thr	Ala	Gly	Ser	Leu	Gln	Gly
		450				455					460				
Gln	Trp	Arg	Gly	Ala	Ala	Gly	Thr	Ala	Ala	Gln	Ala	Ala	Val	Val	Arg
465					470					475					480
Phe	Gln	Gln	Ala	Ala	Asn	Lys	Gln	Lys	Gln	Gln	Leu	Asp	Glu	Ile	Ser
			485						490					495	
Thr	Asn	Ile	Arg	Gln	Ala	Gly	Val	Gln	Tyr	Ser	Arg	Ala	Asp	Gln	Gln
			500						505				510		
Gln	Gln	Gln	Ala	Leu	Ser	Ser	Gln	Met	Gly	Phe	Val	Pro	Thr	Thr	Ala
		515					520					525			
Ala	Ser	Pro	Pro	Ser	Thr	Ala	Ala	Ala	Pro	Pro	Ala	Pro	Ala	Thr	Pro
		530				535					540				
Val	Ala	Pro	Pro	Pro	Pro	Ala	Ala	Ala	Asn	Thr	Pro	Asn	Ala	Gln	Pro
545					550					555					560
Gly	Asp	Pro	Asn	Ala	Ala	Pro	Pro	Pro	Ala	Asp	Pro	Asn	Ala	Pro	Pro
			565						570					575	
Pro	Pro	Val	Ile	Ala	Pro	Asn	Ala	Pro	Gln	Pro	Val	Arg	Ile	Asp	Asn
		580						585						590	
Pro	Val	Gly	Gly	Phe	Ser	Phe	Ala	Leu	Pro	Ala	Gly	Trp	Val	Glu	Ser
		595					600					605			
Asp	Ala	Ala	His	Phe	Asp	Tyr	Gly	Ser	Ala	Leu	Leu	Ser	Lys	Thr	Thr
		610				615						620			

Gly	Asp	Pro	Pro	Phe	Pro	Gly	Gln	Pro	Pro	Pro	Val	Ala	Asn	Asp	Thr	625	630	635	640
Arg	Ile	Val	Leu	Gly	Arg	Leu	Asp	Gln	Lys	Leu	Tyr	Ala	Ser	Ala	Glu	645	650	655	
Ala	Thr	Asp	Ser	Lys	Ala	Ala	Ala	Arg	Leu	Gly	Ser	Asp	Met	Gly	Gln	660	665	670	
Phe	Tyr	Met	Pro	Tyr	Pro	Gly	Thr	Arg	Ile	Asn	Gln	Gln	Thr	Val	Ser	675	680	685	
Leu	Asp	Ala	Asn	Gly	Val	Ser	Gly	Ser	Ala	Ser	Tyr	Tyr	Glu	Val	Lys	690	695	700	
Phe	Ser	Asp	Pro	Ser	Lys	Pro	Asn	Gly	Gln	Ile	Trp	Thr	Gly	Val	Ile	705	710	715	720
Gly	Ser	Pro	Ala	Ala	Asn	Ala	Pro	Asp	Ala	Gly	Pro	Pro	Gln	Arg	Trp	725	730	735	
Phe	Val	Val	Trp	Leu	Gly	Thr	Ala	Asn	Asn	Pro	Val	Asp	Lys	Gly	Ala	740	745	750	
Ala	Lys	Ala	Leu	Ala	Glu	Ser	Ile	Arg	Pro	Leu	Val	Ala	Pro	Pro	Pro	755	760	765	
Ala	Pro	Ala	Pro	Ala	Pro	Ala	Glu	Pro	Ala	Pro	Ala	Pro	Ala	Pro	Ala	770	775	780	
Gly	Glu	Val	Ala	Pro	Thr	Pro	Thr	Thr	Pro	Thr	Pro	Gln	Arg	Thr	Leu	785	790	795	800
Pro	Ala																		

## CLAIMS

We claim:

1. A polypeptide comprising an antigenic portion of a soluble *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen has an N-terminal sequence selected from the group consisting of:

- (a) Asp-Pro-Val-Asp-Ala-Val-Ile-Asn-Thr-Thr-Cys-Asn-Tyr-Gly-Gln-Val-Val-Ala-Ala-Leu (SEQ ID NO: 115);
- (b) Ala-Val-Glu-Ser-Gly-Met-Leu-Ala-Leu-Gly-Thr-Pro-Ala-Pro-Ser (SEQ ID NO: 116);
- (c) Ala-Ala-Met-Lys-Pro-Arg-Thr-Gly-Asp-Gly-Pro-Leu-Glu-Ala-Ala-Lys-Glu-Gly-Arg (SEQ ID NO: 117);
- (d) Tyr-Tyr-Trp-Cys-Pro-Gly-Gln-Pro-Phe-Asp-Pro-Ala-Trp-Gly-Pro (SEQ ID NO: 118);
- (e) Asp-Ile-Gly-Ser-Glu-Ser-Thr-Glu-Asp-Gln-Gln-Xaa-Ala-Val (SEQ ID NO: 119);
- (f) Ala-Glu-Glu-Ser-Ile-Ser-Thr-Xaa-Glu-Xaa-Ile-Val-Pro (SEQ ID NO: 120);
- (g) Asp-Pro-Glu-Pro-Ala-Pro-Pro-Val-Pro-Thr-Thr-Ala-Ala-Ser-Pro-Pro-Ser (SEQ ID NO: 121);
- (h) Ala-Pro-Lys-Thr-Tyr-Xaa-Glu-Glu-Leu-Lys-Gly-Thr-Asp-Thr-Gly (SEQ ID NO: 122);
- (i) Asp-Pro-Ala-Ser-Ala-Pro-Asp-Val-Pro-Thr-Ala-Ala-Gln-Leu-Thr-Ser-Leu-Leu-Asn-Ser-Leu-Ala-Asp-Pro-Asn-Val-Ser-Phe-Ala-Asn (SEQ ID NO: 123); and
- (j) Ala-Pro-Glu-Ser-Gly-Ala-Gly-Leu-Gly-Gly-Thr-Val-Gln-Ala-Gly; (SEQ ID NO: 131)

wherein Xaa may be any amino acid.

2. A polypeptide comprising an immunogenic portion of an *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen has an N-terminal sequence selected from the group consisting of:

- (a) Asp-Pro-Pro-Asp-Pro-His-Gln-Xaa-Asp-Met-Thr-Lys-Gly-Tyr-Tyr-Pro-Gly-Gly-Arg-Arg-Xaa-Phe; (SEQ ID NO: 124) and
- (b) Xaa-Tyr-Ile-Ala-Tyr-Xaa-Thr-Thr-Ala-Gly-Ile-Val-Pro-Gly-Lys-Ile-Asn-Val-His-Leu-Val; (SEQ ID NO: 132), wherein Xaa may be any amino acid.

3. A polypeptide comprising an antigenic portion of a soluble *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen comprises an amino acid sequence encoded by a DNA sequence selected from the group consisting of the sequences recited in SEQ ID NOS: 1, 2, 4-10, 13-25, 52, 94 and 96, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID NOS: 1, 2, 4-10, 13-25, 52, 94 and 96 or a complement thereof under moderately stringent conditions.

4. A polypeptide comprising an antigenic portion of a *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen comprises an amino acid sequence encoded by a DNA sequence selected from the group consisting of the sequences recited in SEQ ID NOS: 26-51, 133, 134, 158-178 and 196, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID NOS: 26-51, 133, 134, 158-178 and 196 or a complement thereof under moderately stringent conditions.

5. A DNA molecule comprising a nucleotide sequence encoding a polypeptide according to any one of claims 1-4.

6. A recombinant expression vector comprising a DNA molecule according to claim 5.

7. A host cell transformed with an expression vector according to claim 6.

8. The host cell of claim 7 wherein the host cell is selected from the group consisting of *E. coli*, yeast and mammalian cells.

9. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:

(a) contacting a biological sample with one or more polypeptides according to any of claims 1-4; and

(b) detecting in the sample the presence of antibodies that bind to at least one of the polypeptides, thereby detecting *M. tuberculosis* infection in the biological sample.

10. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:

(a) contacting a biological sample with a polypeptide having an N-terminal sequence selected from the group consisting of sequences provided in SEQ ID NO: 129 and 130; and

(b) detecting in the sample the presence of antibodies that bind to at least one of the polypeptides, thereby detecting *M. tuberculosis* infection in the biological sample.

11. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:

(a) contacting a biological sample with one or more polypeptides encoded by a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195 and 198, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195 and 198; and

(b) detecting in the sample the presence of antibodies that bind to at least one of the polypeptides, thereby detecting *M. tuberculosis* infection in the biological sample.

12. The method of any one of claims 9-11 wherein step (a) additionally comprises contacting the biological sample with a 38 kD *M. tuberculosis* antigen and step (b) additionally comprises detecting in the sample the presence of antibodies that bind to the 38 kD *M. tuberculosis* antigen.

13. The method of any one of claims 9-11 wherein the polypeptide(s) are bound to a solid support.

14. The method of claim 13 wherein the solid support comprises nitrocellulose, latex or a plastic material.

15. The method of any one of claims 9-11 wherein the biological sample is selected from the group consisting of whole blood, serum, plasma, saliva, cerebrospinal fluid and urine.

16. The method of claim 15 wherein the biological sample is whole blood or serum.

17. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:

(a) contacting the sample with at least two oligonucleotide primers in a polymerase chain reaction, wherein at least one of the oligonucleotide primers is specific for a DNA molecule according to claim 5; and

(b) detecting in the sample a DNA sequence that amplifies in the presence of the oligonucleotide primers, thereby detecting *M. tuberculosis* infection.

18. The method of claim 17, wherein at least one of the oligonucleotide primers comprises at least about 10 contiguous nucleotides of a DNA molecule according to claim 5.

19. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:

(a) contacting the sample with at least two oligonucleotide primers in a polymerase chain reaction, wherein at least one of the oligonucleotide primers is specific for a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195 and 198; and

(b) detecting in the sample a DNA sequence that amplifies in the presence of the first and second oligonucleotide primers, thereby detecting *M. tuberculosis* infection.

20. The method of claim 19, wherein at least one of the oligonucleotide primers comprises at least about 10 contiguous nucleotides of a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195 and 198.

21. The method of claims 17 or 19 wherein the biological sample is selected from the group consisting of whole blood, sputum, serum, plasma, saliva, cerebrospinal fluid and urine.

22. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:

(a) contacting the sample with one or more oligonucleotide probes specific for a DNA molecule according to claim 5; and

(b) detecting in the sample a DNA sequence that hybridizes to the oligonucleotide probe, thereby detecting *M. tuberculosis* infection.



23. The method of claim 22 wherein the probe comprises at least about 15 contiguous nucleotides of a DNA molecule according to claim 5.

24. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:

(a) contacting the sample with one or more oligonucleotide probes specific for a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195 and 198; and

(b) detecting in the sample a DNA sequence that hybridizes to the oligonucleotide probe, thereby detecting *M. tuberculosis* infection.

25. The method of claim 24 wherein the oligonucleotide probe comprises at least about 15 contiguous nucleotides of a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195 and 198.

26. The method of claims 22 or 24 wherein the biological sample is selected from the group consisting of whole blood, sputum, serum, plasma, saliva, cerebrospinal fluid and urine.

27. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:

(a) contacting the biological sample with a binding agent which is capable of binding to a polypeptide according to any one of claims 1-4; and

(b) detecting in the sample a protein or polypeptide that binds to the binding agent, thereby detecting *M. tuberculosis* infection in the biological sample.

28. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:

- (a) contacting the biological sample with a binding agent which is capable of binding to a polypeptide having an N-terminal sequence selected from the group consisting of sequences provided in SEQ ID NO: 129 and 130; and
- (b) detecting in the sample a protein or polypeptide that binds to the binding agent, thereby detecting *M. tuberculosis* infection in the biological sample.

29. A method for detecting *M. tuberculosis* infection in a biological sample, comprising:

- (a) contacting the biological sample with a binding agent which is capable of binding to a polypeptide encoded by a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195 and 198, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195 and 198; and
- (b) detecting in the sample a protein or polypeptide that binds to the binding agent, thereby detecting *M. tuberculosis* infection in the biological sample.

30. The method of any one of claims 27-29 wherein the binding agent is a monoclonal antibody.

31. The method of any one of claims 27-29 wherein the binding agent is a polyclonal antibody.

32. A diagnostic kit comprising:

- (a) one or more polypeptides according to any of claims 1-4; and
- (b) a detection reagent.

33. A diagnostic kit comprising:

- (a) one or more polypeptides having an N-terminal sequence selected from the group consisting of sequences provided in SEQ ID NO: 129 and 130; and
- (b) a detection reagent.

34. A diagnostic kit comprising:
- (a) one or more polypeptides encoded by a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195 and 198, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195 and 198; and
  - (b) a detection reagent.
35. The kit of any one of claims 32-34 wherein the polypeptide(s) are immobilized on a solid support.
36. The kit of claim 35 wherein the solid support comprises nitrocellulose, latex or a plastic material.
37. The kit of any one of claims 32-34 wherein the detection reagent comprises a reporter group conjugated to a binding agent.
38. The kit of claim 37 wherein the binding agent is selected from the group consisting of anti-immunoglobulins, Protein G, Protein A and lectins.
39. The kit of claim 37 wherein the reporter group is selected from the group consisting of radioisotopes, fluorescent groups, luminescent groups, enzymes, biotin and dye particles.
40. A diagnostic kit comprising at least two oligonucleotide primers, at least one of the oligonucleotide primers being specific for a DNA molecule according to claim 5.

41. A diagnostic kit according to claim 40, wherein at least one of the oligonucleotide primers comprises at least about 10 contiguous nucleotide of a DNA molecule according to claim 5.

42. A diagnostic kit comprising a at least two oligonucleotide primers, at least one of the primers being specific for a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195 and 198.

43. A diagnostic kit according to claim 42, wherein at least one of the oligonucleotide primers comprises at least about 10 contiguous nucleotide of a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195 and 198.

44. A diagnostic kit comprising at least one oligonucleotide probe, the oligonucleotide probe being specific for a DNA molecule according to claim 5.

45. A kit according to claim 44, wherein the oligonucleotide probe comprises at least about 15 contiguous nucleotides of a DNA molecule according to claim 5.

46. A diagnostic kit comprising at least one oligonucleotide probe, the oligonucleotide probe being specific for a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195 and 198.

47. A kit according to claim 46, wherein the oligonucleotide probe comprises at least about 15 contiguous nucleotides of a DNA sequence selected from the group consisting of SEQ ID NOS: 3, 11, 12, 135, 136, 151-155, 184-188, 194-195 and 198.

48. A monoclonal antibody that binds to a polypeptide according to any of claims 1-4.

49. A polyclonal antibody that binds to a polypeptide according to any of claims 1-4.

50. A fusion protein comprising two or more polypeptides according to any one of claims 1-4.

51. A fusion protein comprising one or more polypeptides according to any one of claims 1-4 and ESAT-6 (SEQ ID NO: 99).

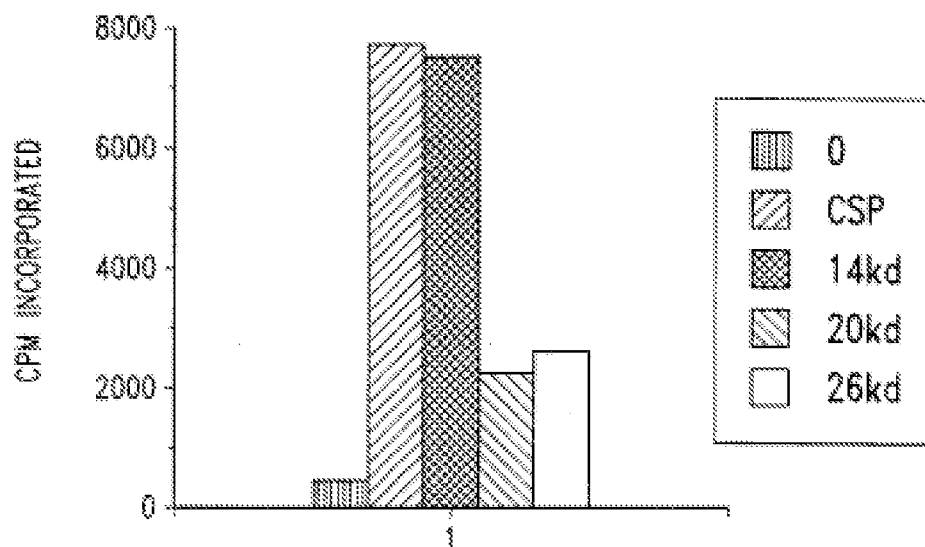
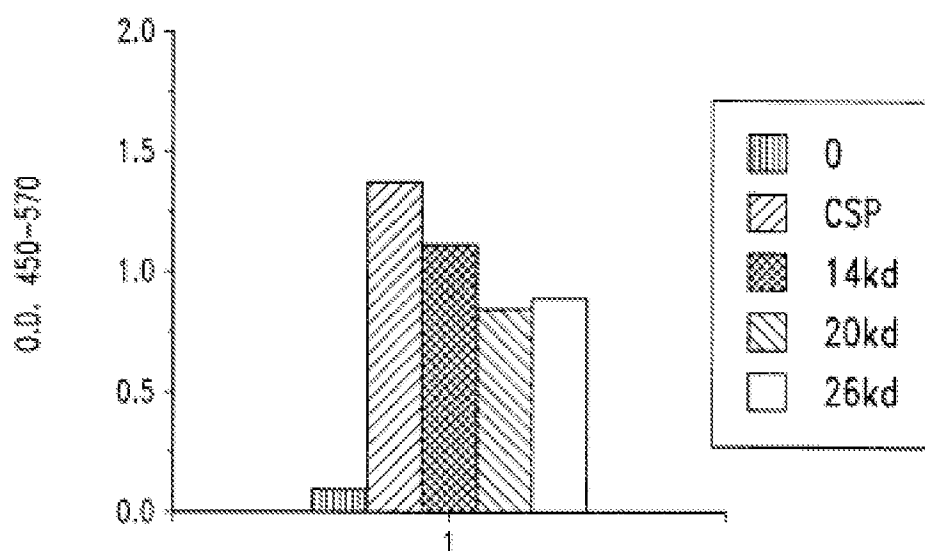
52. A fusion protein comprising a polypeptide having an N-terminal sequence selected from the group of sequences provided in SEQ ID NOS: 129 and 130.

53. A fusion protein comprising one or more polypeptides according to any one of claims 1-4 and the *M. tuberculosis* antigen 38 kD (SEQ ID NO: 150).

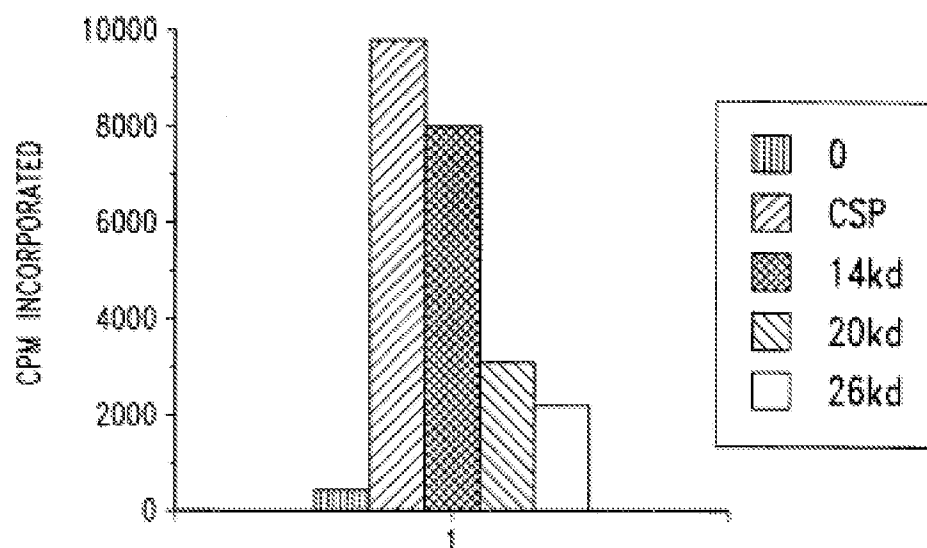
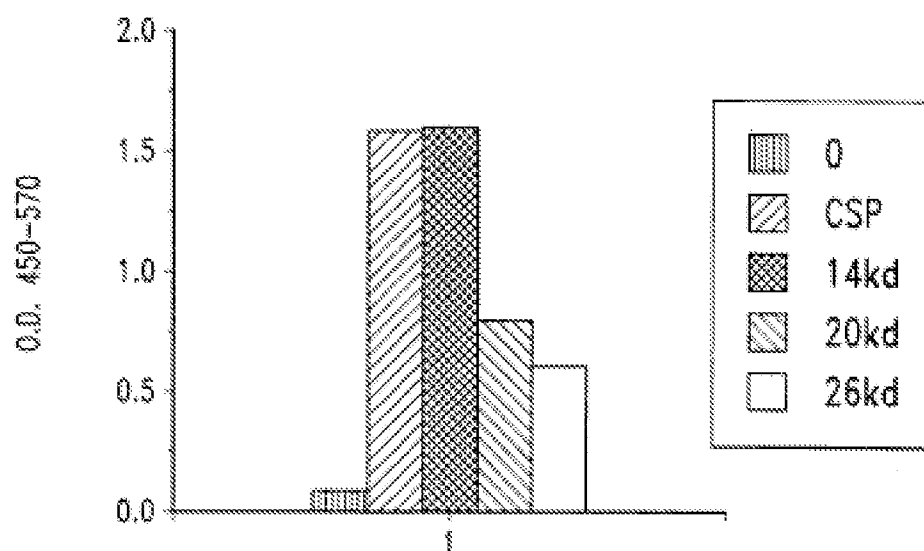
54. A diagnostic kit comprising:

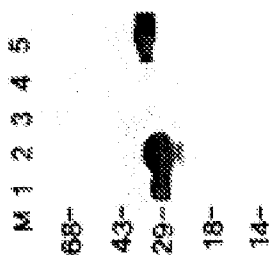
- (a) one or more fusion proteins according to any one of claims 50-53; and
- (b) a detection reagent.

1/13

*Fig. 1A-1**Fig. 1A-2*

2/13

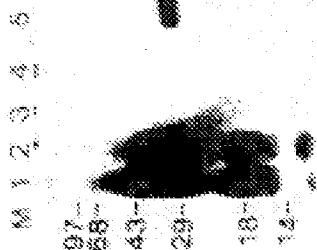
*Fig. 1B-1**Fig. 1B-2*



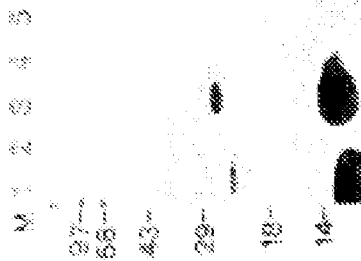
I  
II  
Fig. 2B



I  
II  
Fig. 2D



I  
II  
Fig. 2A



I  
II  
Fig. 2C



4/13

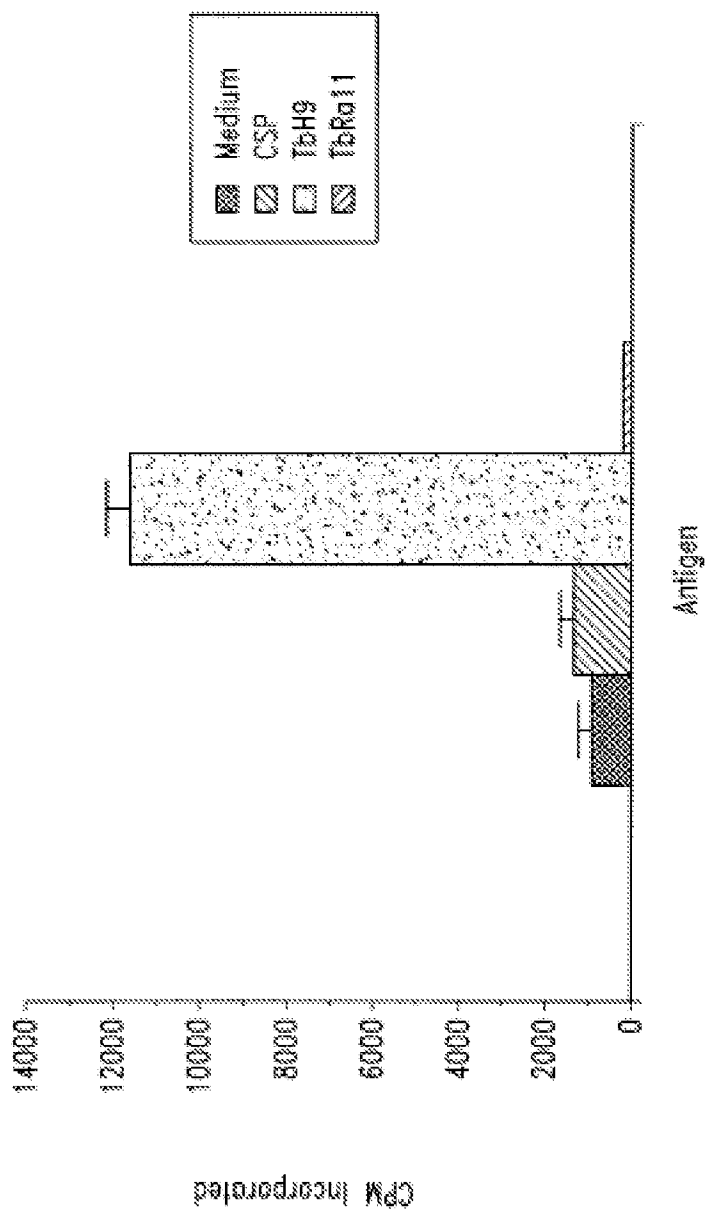
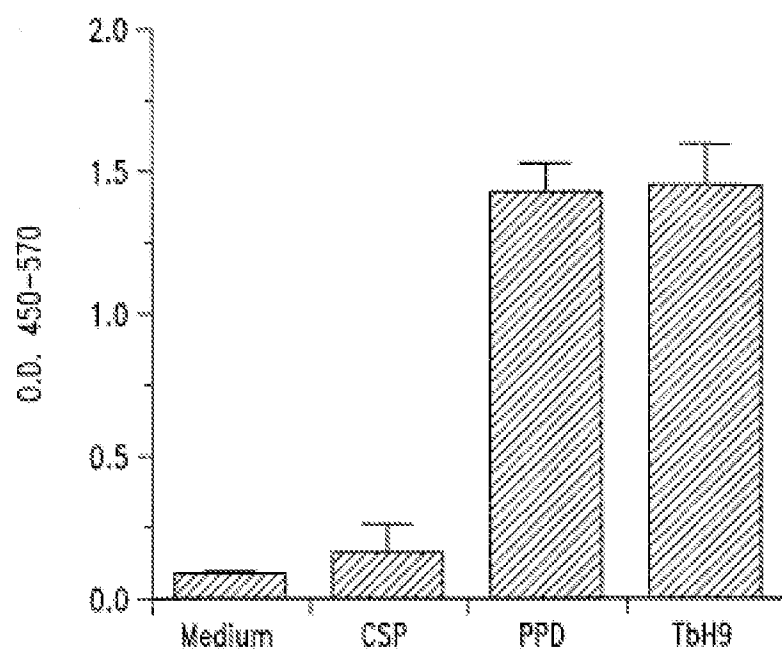


Fig. 3A

5/13

*Fig. 3B*

6/13

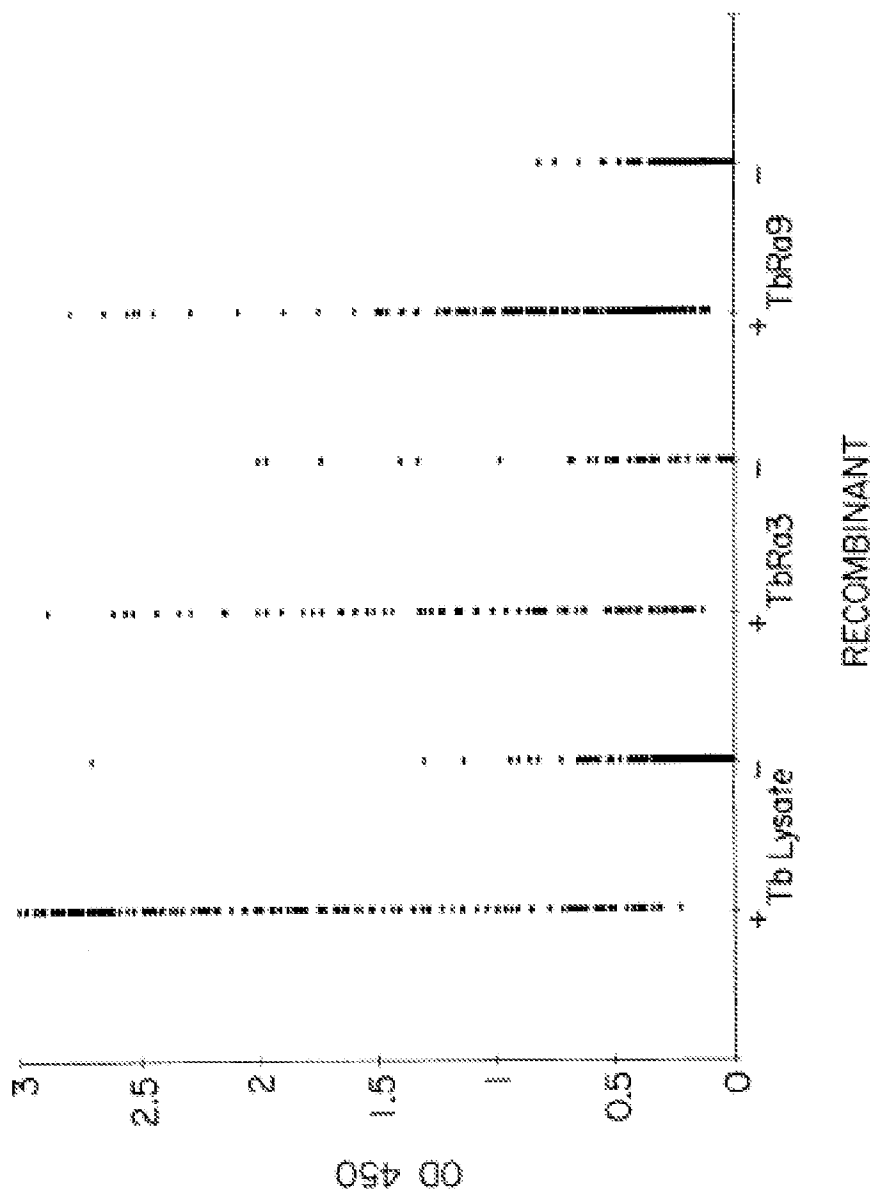


Fig. 4

7/13

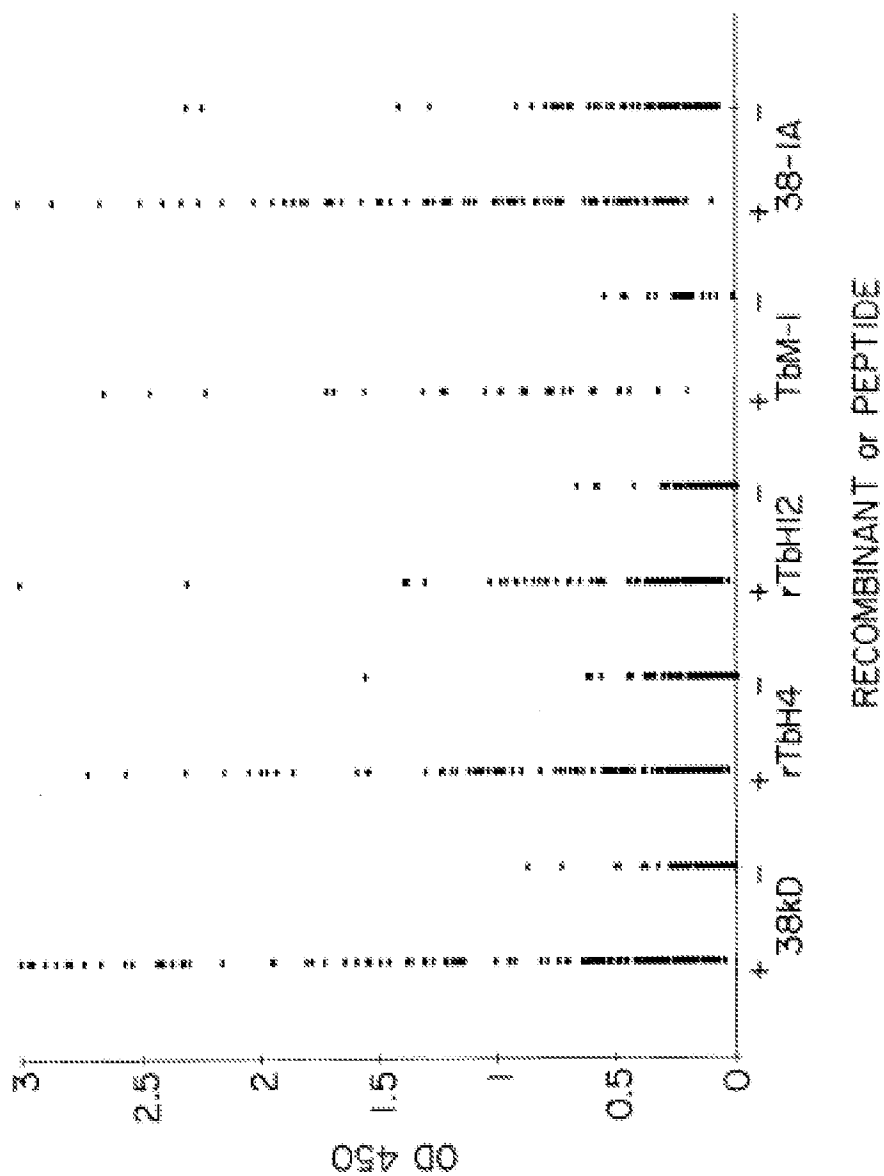


Fig. 5

8/13

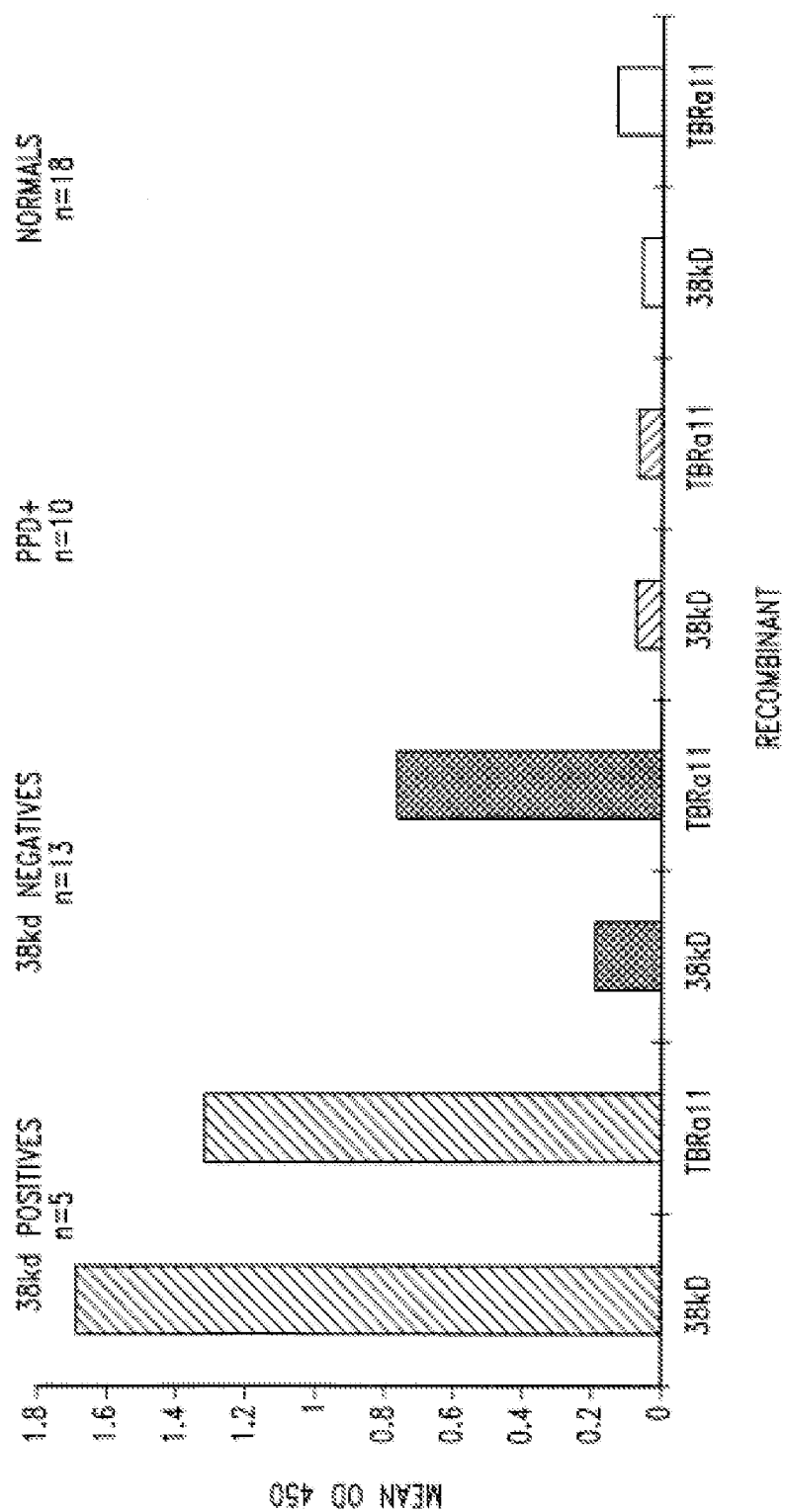


Fig. 6

9/13

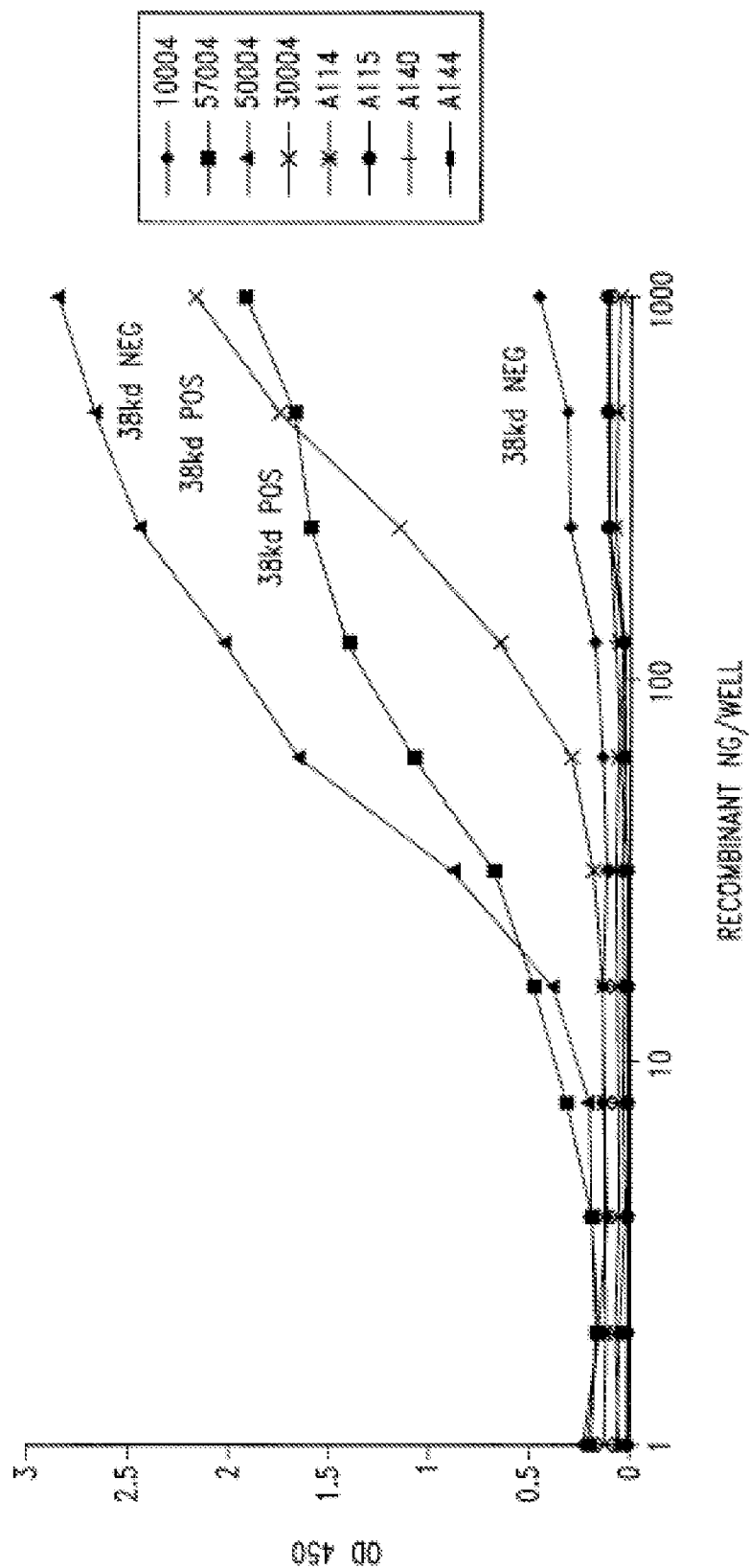


Fig. 7

10/13

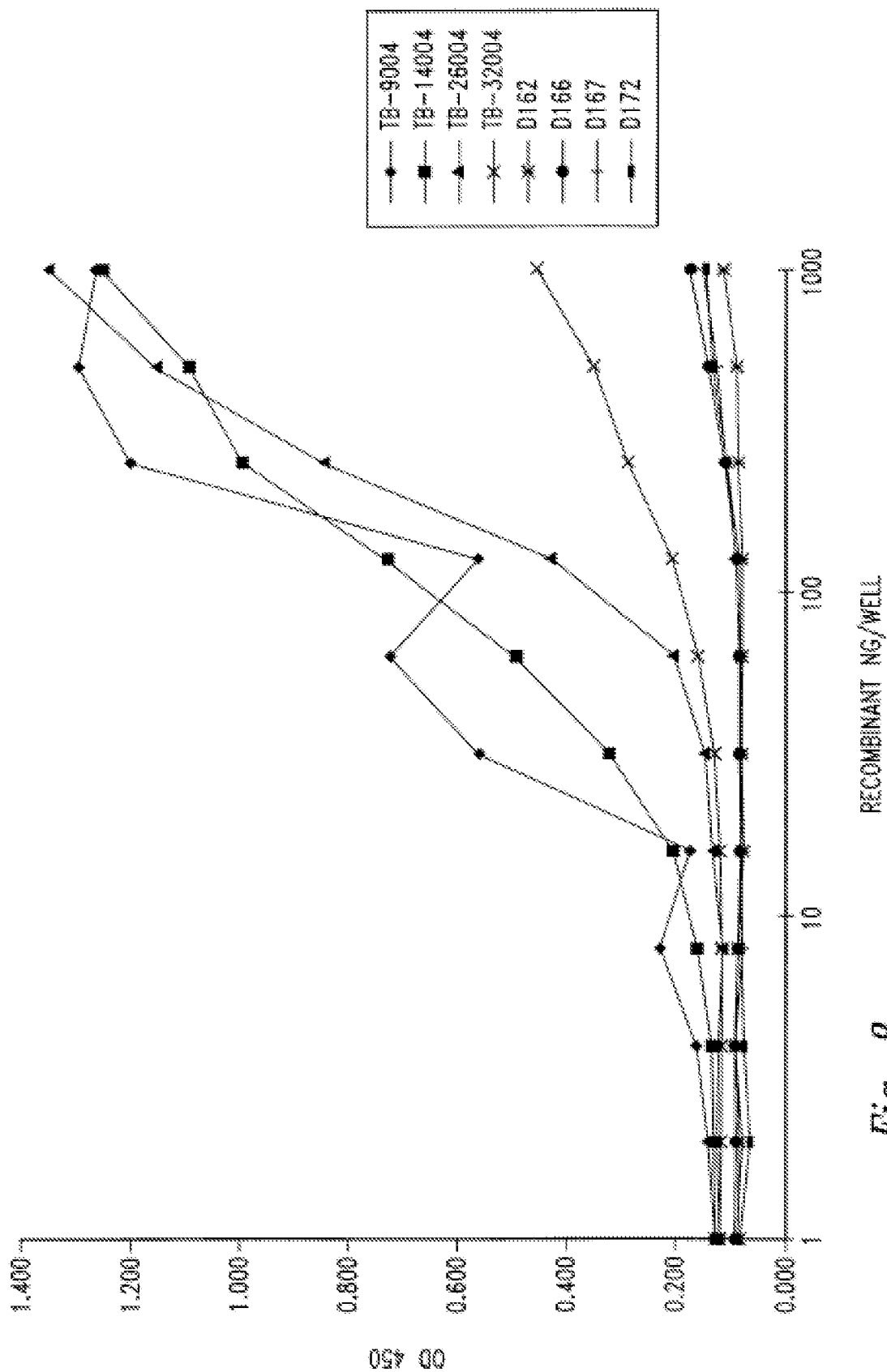


Fig. 8

11/13

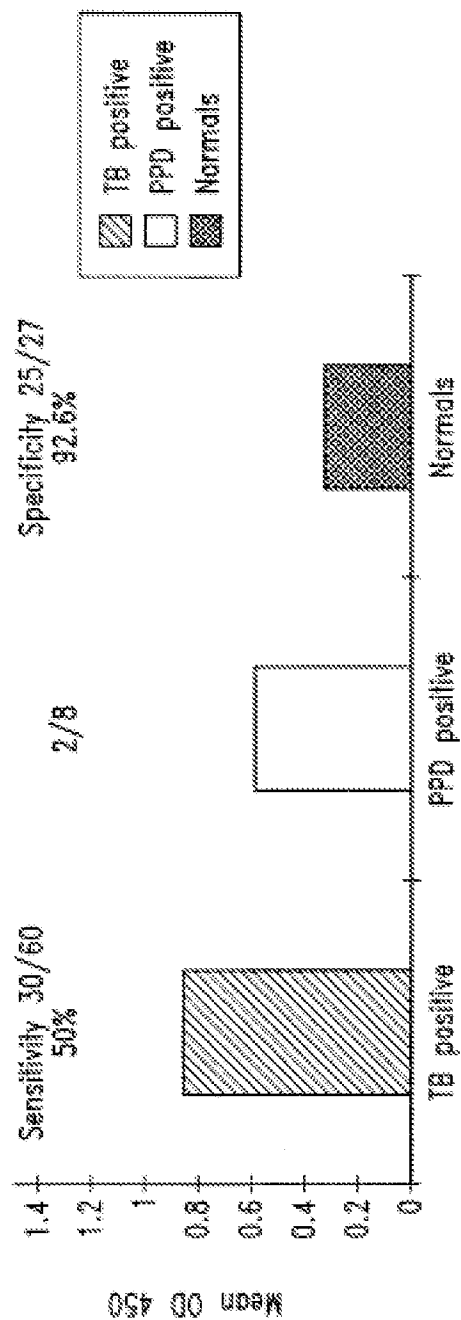


Fig. 9



12/13

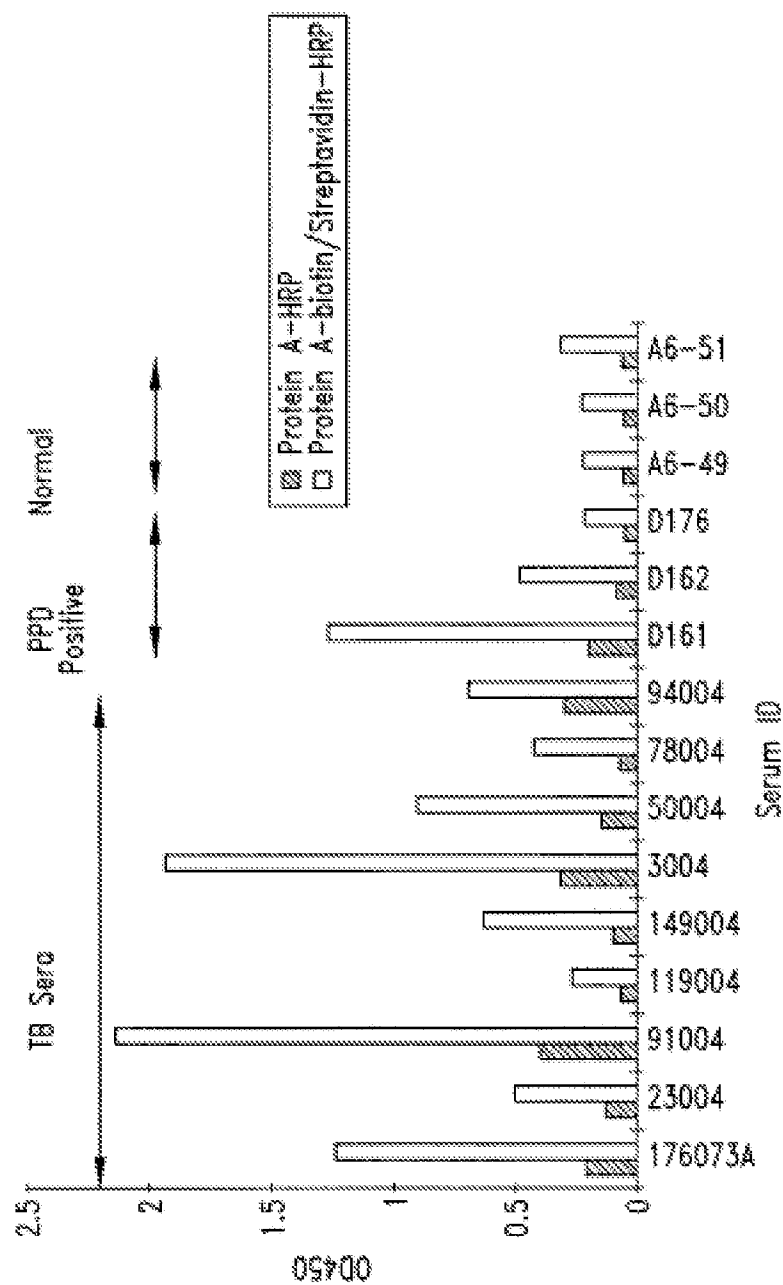


Fig. 10

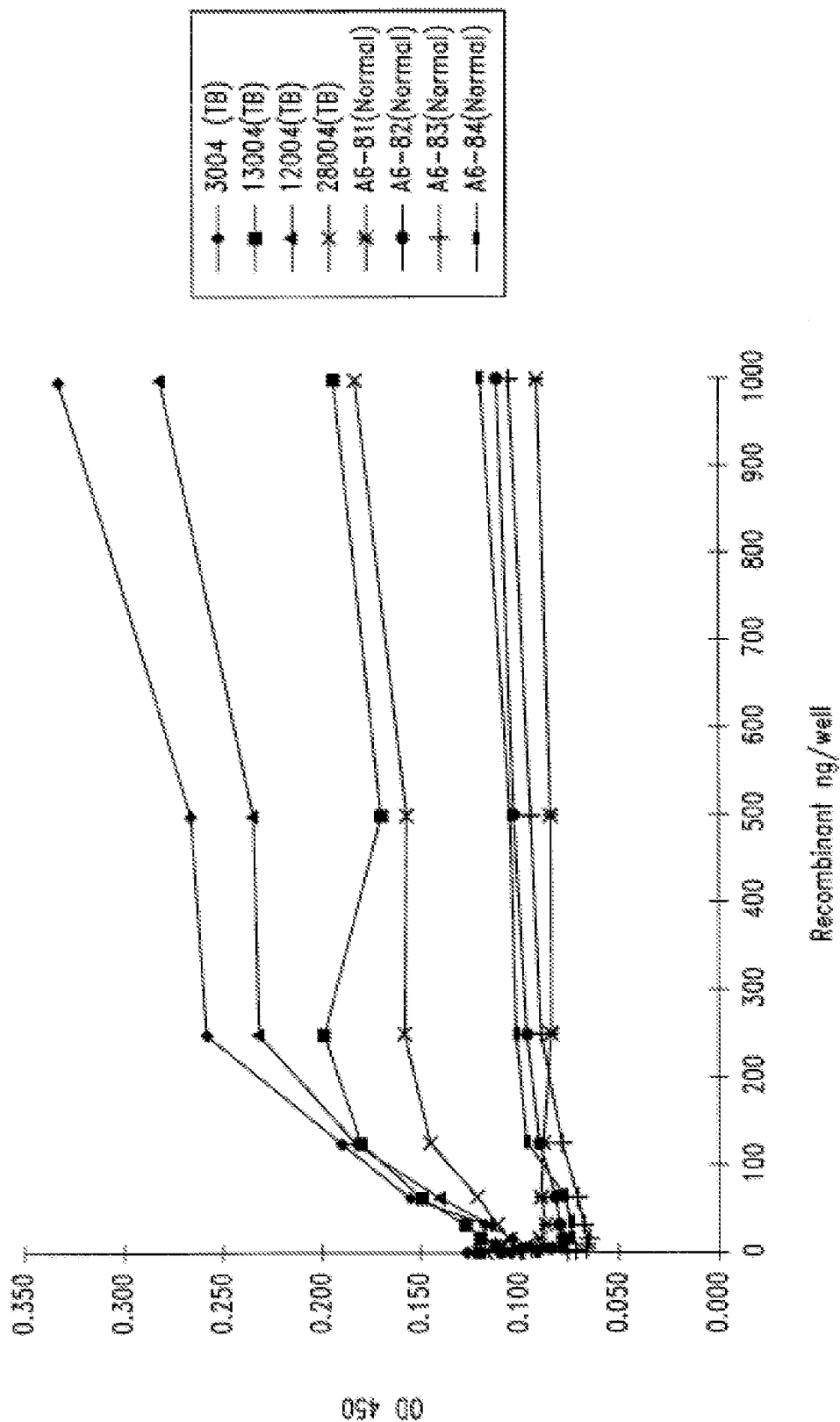


Fig. 11